

Analyse de données de fréquentation du Parc national des Écrins



Victor Cazalis (Conservara)

Mai à Novembre 2025

Destinataire :

Parc national des Ecrins

Domaine de Charance

05000 Gap



Projet BiodivTourAlps n°20140

Table des matières

Résumé.....	3
Remerciements.....	3
Contexte.....	4
Synthèse des résultats.....	5
Résumé des recommandations.....	11
Préparation des données.....	13
Préparation des données internes.....	14
Maisons de parc.....	17
Refuges.....	17
Eco-compteurs.....	17
Pièges photo.....	18
Enquêtes.....	18
Compteurs routiers.....	18
Préparation des données externes.....	20
Outdoorvision.....	20
Strava.....	22
iNaturalist.....	22
PlantNet.....	23
Objectif 1 : cartographier la fréquentation à l'échelle du parc.....	24
Cartographie interactive du PNE.....	25
Cartographie standardisée du PNE.....	27
Méthodes.....	27
Résultats.....	30
Discussion.....	33
Cartographie des zones blanches du suivi interne.....	36
Méthodes.....	36
Résultats.....	39
Discussion.....	43
Analyse temporelle de la fréquentation.....	45
Tendance inter-annuelle.....	45
Analyse saisonnière.....	46
Discussion.....	48
Objectif 2 : étudier la fréquentation spécifique du vallon du Lauvitel.....	50
Diffusion spatiale dans le vallon.....	51
Distribution temporelle fine.....	54
Répartition des différentes activités.....	55
Autres questions.....	57
Objectif 3 : utiliser les données de fréquentation pour la protection de la biodiversité.....	59
Données de biodiversité sensible.....	59
Sites de nidification d'aigle royal.....	59
Stations botaniques de chardon bleu.....	59
Habitats prioritaires Natura2000.....	60
Zones humides.....	60
Zones de quiétude pour tétras.....	60
Méthodes.....	60
Résultats.....	61
Discussion.....	64
Discussion générale.....	66
Livrables supplémentaires.....	68
Références.....	69
Annexes.....	70

Résumé

Dans le cadre du projet BiodivTourAlps, le Parc national des Ecrins s'intéresse à la fréquentation humaine de son territoire par des activités de randonnée, course ou cyclisme. Le parc dispose de plusieurs dispositifs internes collectant des données de fréquentation (compteurs pédestres, compteurs routiers, pièges photo placés sur les chemins de randonnée, comptage des visites de maisons de parc, comptage des nuitées en refuge, enquêtes réalisées sur les chemins de randonnée). A ces données internes s'ajoutent des données spatiales auxquelles le parc a accès, soit sportives (Strava et Outdoorvision qui collectent les traces GPS de randonneurs et coureurs), soit d'activités naturalistes (avec iNaturalist et PlantNet qui permettent de cartographier les lieux de fréquentation des naturalistes amateurs).

La mission d'analyse présentée dans ce rapport visait à explorer l'utilisation de ces données pour voir à quelles questions elles permettent de répondre et comment elles peuvent être mieux valorisées. Elle a permis de développer une méthode de préparation des différentes données internes et externes pour les rendre interopérables, en proposant des améliorations sur le formatage et la gestion des données internes pour une prise en main plus rapide. Une fois prêtes à l'emploi, ces données ont été intégrées dans une carte interactive, qui permet de les visualiser et ainsi d'offrir un aperçu de l'ensemble des données de fréquentation disponibles dans le parc.

Les différentes analyses ont visé à cartographier de manière standardisée la fréquentation du parc, à identifier les zones où les dispositifs de suivi étaient insuffisants, à étudier les variations temporelles de la fréquentation (tendances pluriannuelles et effets saisonniers), à étudier la dispersion de la fréquentation à une échelle fine dans le vallon du Lauvitel, et à étudier l'interaction entre fréquentation humaine et biodiversité sensible. Les résultats de ces analyses sont résumés dans la synthèse ci-dessous et détaillés dans ce rapport.

A l'issue de ces analyses, 13 recommandations ont été rédigées, visant à améliorer la gestion des données de fréquentation, à collecter des données manquantes, et à développer de nouvelles méthodes d'analyse. Elles sont résumées à la fin de ce rapport.

Remerciements

Je tiens à remercier l'équipe du Parc national des Ecrins de m'avoir confié cette mission très intéressante et pour leur accompagnement au cours de l'étude. Merci en particulier à Juliette pour les échanges réguliers et l'aide à la compréhension des données. Merci également à Yoann et Pierrick pour leurs retours pertinents et bienveillants lors des réunions d'étape qui m'ont permis d'avancer dans la bonne direction. Enfin, merci aux contributeurs et contributrices des données Strava, Outdoorvision, iNaturalist et PlantNet, grâce à qui cette mission a pu être largement enrichie, et merci aux personnes développant le logiciel R et ses multiples packages utilisés dans cette analyse.

Contexte

Le projet BiodivTourAlps « Biodiversité et tourisme : synergie entre Parcs alpins », a pour objectif général d'agir sur les effets des activités touristiques sur les milieux et les ressources naturelles au sein des espaces protégés transfrontaliers, pour répondre aux enjeux de préservation de la biodiversité. Il a été lancé en octobre 2023, pour une durée de 3 ans et regroupe sept partenaires français et italiens : les Parcs nationaux des Écrins, de la Vanoise, du Mercantour (chef de file), le Parc national du Gran Paradiso, le Parco naturale regionale delle Alpi Li-guri, l'Ente di gestione delle Aree Protette delle Alpi Marittime et l'Ente di gestione dei Parchi delle AlpiCozie.

Les actions s'articulent autour des thématiques de la fréquentation et des interactions humains-nature. Une de ces actions est la mise en place de méthodologies et d'une boîte à outils pour faciliter le suivi et la gestion des flux. Le projet prévoit également le déploiement de protocoles de suivi pour l'acquisition de nouvelles données de fréquentation. C'est dans le cadre de cette action que cette prestation a été réalisée.

Le Parc national des Écrins (PNE) utilise des dispositifs de suivi de sa fréquentation depuis de nombreuses années, ce qui a permis la collecte de données à différentes échelles spatiales et temporelles, utilisant différents protocoles, et se focalisant sur différents types d'activités. L'objectif de cette prestation était de mettre en place des protocoles d'analyse de ces données pour les valoriser au mieux, tout en identifiant les questions auxquelles ces données peuvent répondre et quelles données supplémentaires permettraient de répondre à plus de questions. L'objectif n'est pas de donner des réponses définitives à une ou deux grandes questions, mais plutôt d'explorer, dans différentes directions, les questions qui peuvent être étudiées avec les données disponibles et de faire avancer les réflexions sur le suivi de la fréquentation dans le PNE.

Cette prestation s'est focalisée sur trois objectifs principaux : (1) la description et cartographie de la fréquentation à l'échelle du PNE, (2) l'étude localisée de la fréquentation d'un site spécifique, (3) l'utilisation des données de fréquentation dans le cadre des missions de protection de la biodiversité.

Synthèse des résultats

Le Parc national des Ecrins (PNE) dispose de nombreux jeux de données lui permettant d'étudier sa fréquentation pour des activités de pleine nature. Plusieurs données internes sont collectées depuis de nombreuses années par des compteurs routiers, des compteurs pédestres (éco-compteurs), des pièges photo, les maisons de parc, les refuges, et par des enquêtes menées sur les chemins de randonnée. Comme cela est détaillé dans le rapport, ces données pourraient être rendues plus faciles à utiliser si elles respectaient un format commun de tableau.

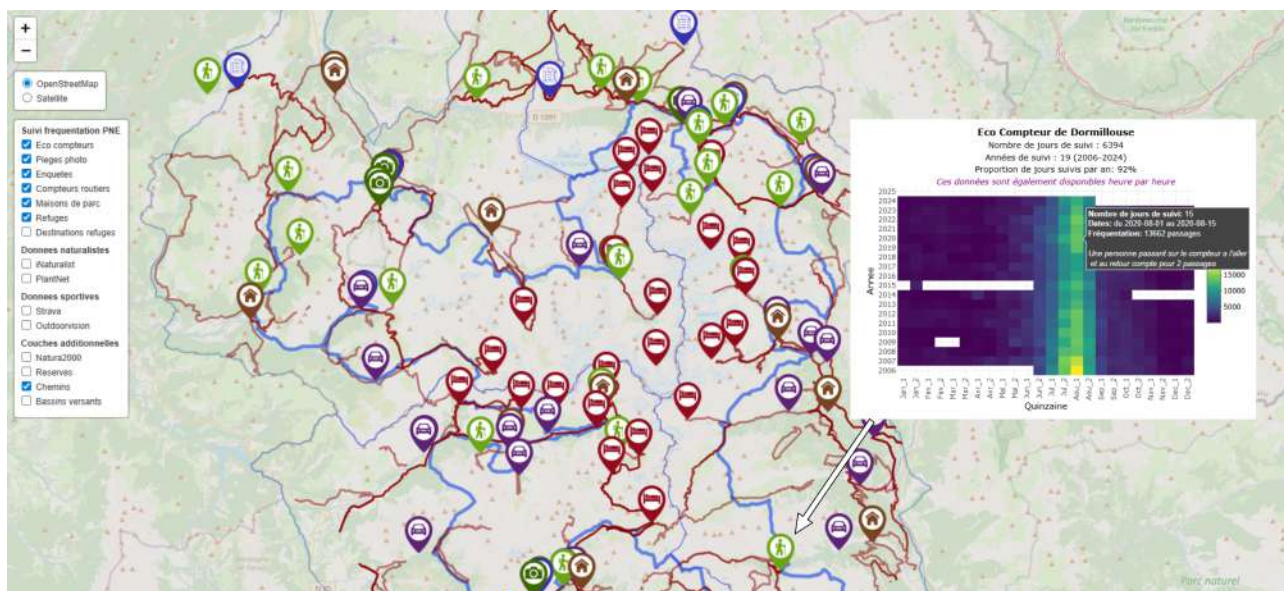
The diagram shows a table with four columns: Nom, Date, Heure, and Comptage. Annotations point to specific features: 'Début tableau en A1' points to the start of the table; 'Inclure les noms dans le tableau + correspondance des noms avec le fichier spatial' points to the 'Nom' column; 'Unique colonne de comptage' points to the 'Comptage' column; 'Aucune cellule fusionnée' points to the 'Comptage' column; 'Fichier unique avec tous les compteurs' points to the 'Nom' column; 'Format date homogène' points to the 'Date' column; 'Pas de ligne "Total" en fin de tableau' points to the bottom of the table; 'Différencier comptage nul (0) dysfonctionnement (NaN)' points to a cell containing 'NaN'.

	A	B	C	D
	Nom	Date	Heure	Comptage
1	La Danchère	01/07/25	10:00:00	2
2	La Danchère	01/07/25	11:00:00	10
3	La Danchère	02/07/25	10:00:00	3
4	La Danchère	02/07/25	11:00:00	11
5	Casset	01/07/25	10:00:00	1
6	Casset	01/07/25	11:00:00	0
7	Casset	02/07/25	10:00:00	NaN
8	Casset	02/07/25	11:00:00	15

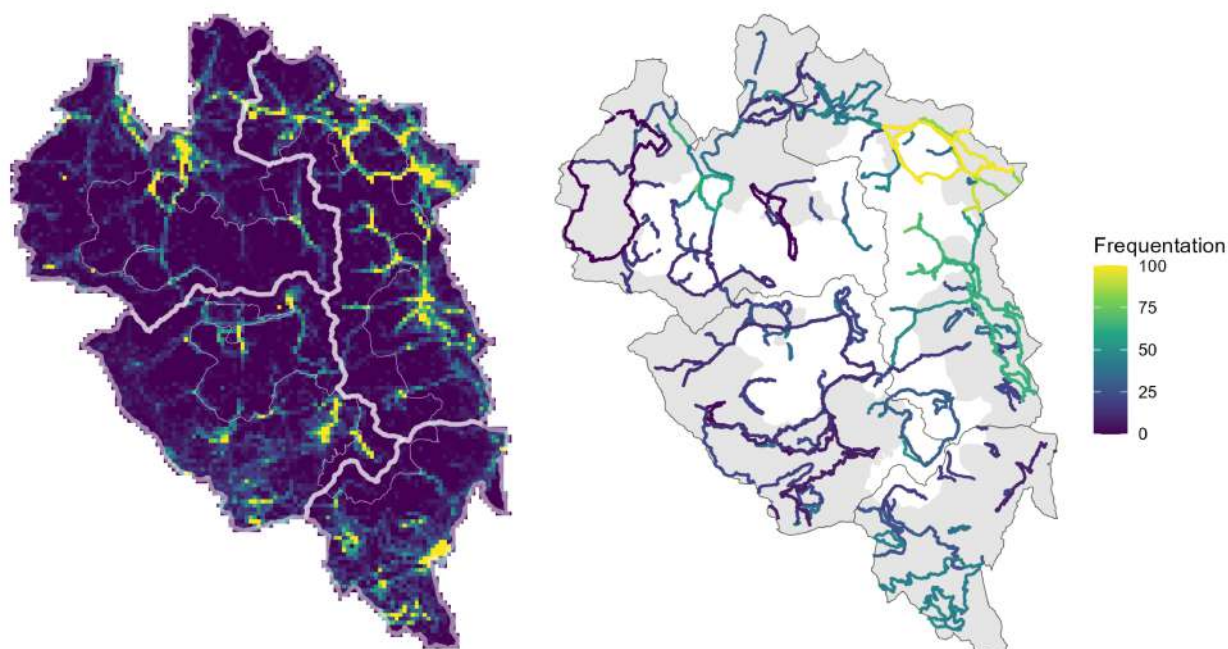
A ces données internes, s'ajoutent des données externes qui sont libres d'accès ou ont été mises à disposition du parc, et permettent d'avoir des informations spatiales sur la fréquentation : Strava, Outdoorvision, iNaturalist et PlantNet. Ces données couvrent l'ensemble du parc, ont une grande finesse spatiale et sont donc un excellent complément aux données internes du parc.

Objectif 1 : cartographier la fréquentation à l'échelle du parc

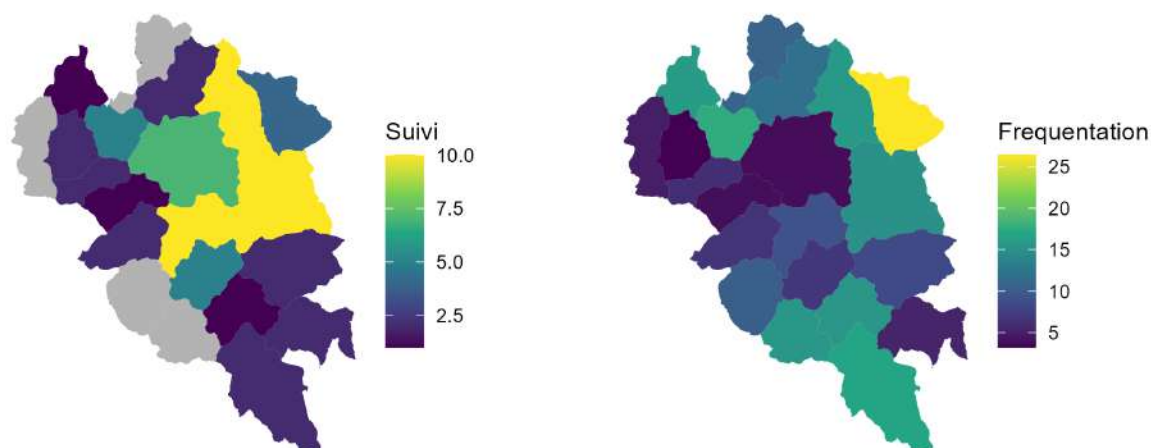
Une première étape pour valoriser l'ensemble de ces données est de pouvoir les visualiser ensemble. Une carte interactive a été créée permettant d'explorer l'ensemble des données internes et externes, afin de voir leur répartition spatiale ainsi que l'intensité de fréquentation qu'elles mesurent. Cette carte interactive peut servir de vitrine des données dont dispose le parc.



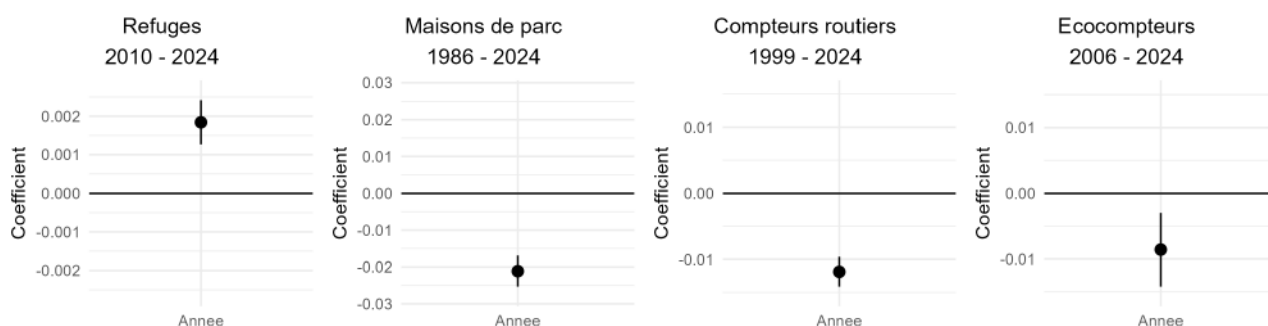
En combinant les données externes qui offrent une cartographie très précise de la fréquentation, et les données d'éco-compteurs qui offrent un comptage exhaustif et précis de certains chemins, deux cartes standardisées de la fréquentation estivale ont été créées. La première est basée sur une grille raster de 500m de résolution, l'autre est basée sur les grands chemins de randonnée du PNE. Ces cartes permettent de constater que les zones autour des Deux Alpes, du Puy Saint Vincent, du Monétier les Bains, d'Orcières et d'Embrun sont particulièrement fréquentées, tout comme les secteurs du Pré de Mme Carle, du Gioberney et du Lauvitel en cœur de parc.



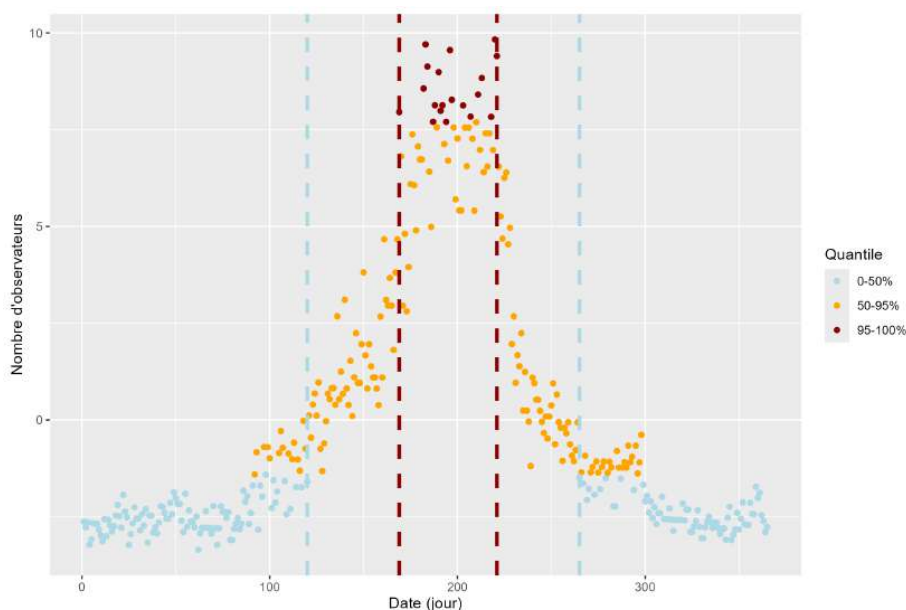
La distribution des dispositifs de suivi mis en place par le PNE est largement concentrée dans le cœur du parc et certains bassins versants ne sont suivis par aucun dispositif de données internes (Clavens, Taillefer, Torrent de la Séveraissette et de la Muande et Pont-du-Fossé/Ancelle) ou par des dispositifs mais aucun éco-compteur (Col d'Ornon à Bourg d'Oisans et Source de la Bonne). Ces bassins versants, notamment les plus fréquentés d'entre eux, devraient être prioritaires pour déployer de nouveaux éco-compteurs ou pièges photo. De même, les éco-compteurs sont principalement déployés dans une tranche médiane d'altitude (1000-2000m) et il serait intéressant d'en placer, au moins sur deux ou trois ans, à d'autres altitudes.



Plusieurs jeux de données mobilisés dans cette analyse permettent une exploration temporelle des données. Les jeux de données de refuge, d'éco-compteurs, de maisons de parc, et de compteurs routiers permettent de calculer une tendance temporelle de la fréquentation sur plus d'une décennie, suggérant une baisse de la fréquentation sauf pour les refuges.

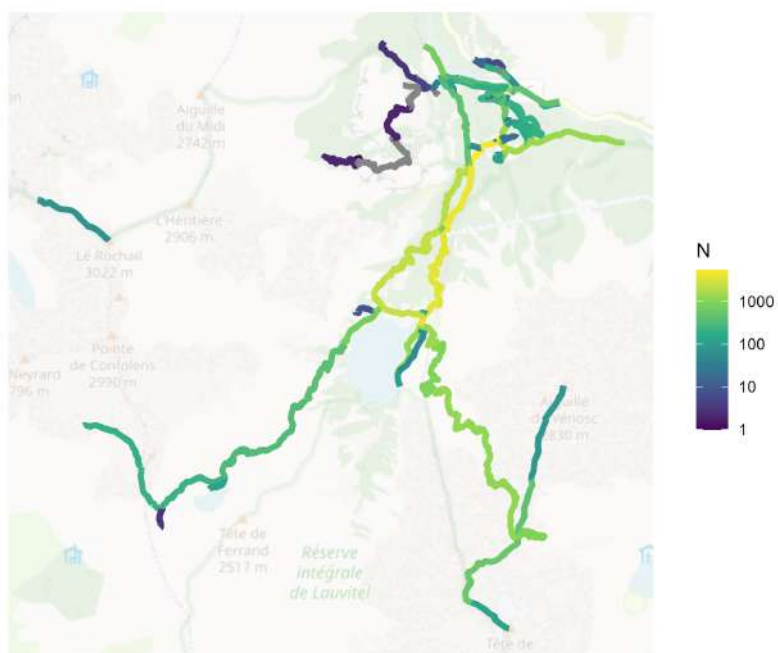


D'autres jeux de données, comme iNaturalist, permettent une analyse fine de la dynamique saisonnière, mettant en évidence un pic de fréquentation entre le 18 Juin et le 9 Août.

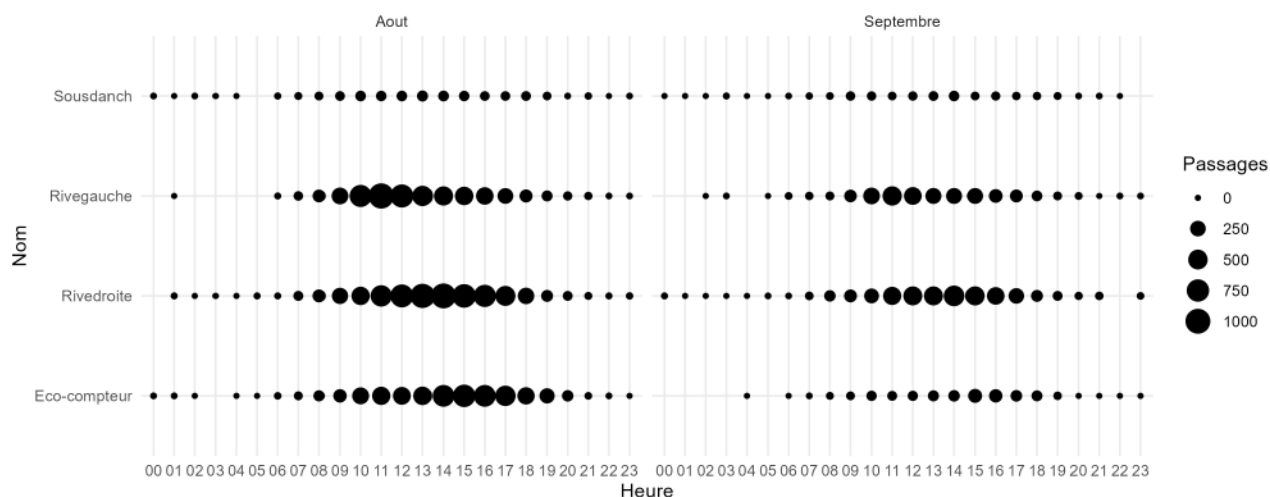


Objectif 2 : étudier la fréquentation spécifique du vallon du Lauvitel

L'étude de la fréquentation peut être plus fine si l'on restreint la zone d'étude, comme par exemple au vallon du Lauvitel. On peut ainsi étudier la diffusion fine entre les différents chemins partant du bas de la vallée (92% arrivant du parking et 8% arrivant des Gourniers). La quasi-totalité des personnes empruntant ce chemin vont jusqu'au lac du Lauvitel, et environ 25% poussent au-delà du lac (20% vers le col du Vallon, 5% vers la Brèche du Périer). Grâce aux données iNaturalist, on peut également observer que les chemins sont globalement bien respectés et que les randonneurs ne s'en écartent qu'aux abords du lac (probablement pour le pique-nique).

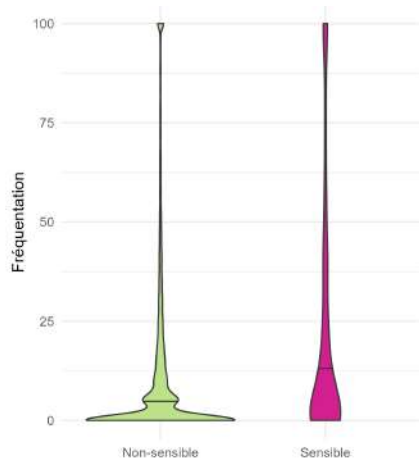
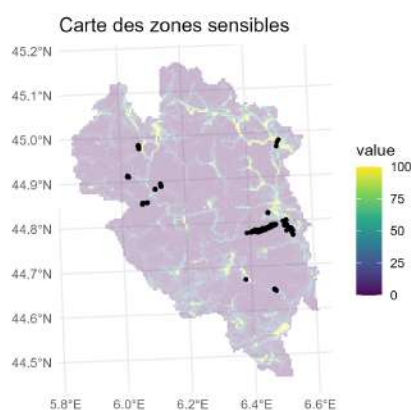


Cette échelle fine permet également d'étudier la distribution par heure, grâce aux données Strava ou pièges photo. La fréquentation estivale du vallon se concentre entre 7h et 17h et les altitudes de plus de 2000m sont vite atteintes (à partir de 8h ou 9h) et voient leur fréquentation diminuer en début d'après-midi. Les pièges photo nous permettent également de voir que la rive gauche est privilégiée à la montée tandis que la rive droite est privilégiée à la descente.



Objectif 3 : utiliser les données de fréquentation pour la protection de la biodiversité

Les cartes de fréquentation standardisée produites dans l'objectif 1 peuvent être utilisées pour identifier les zones de forte interaction entre la fréquentation humaine et la biodiversité sensible. Cela ne signifie pas pour autant que la biodiversité est dérangée par ces activités (il faudrait regarder au cas par cas), mais permet de détecter les zones de forte interface. En testant cinq jeux de données de biodiversité sensible, nous avons trouvé que la fréquentation était significativement moins importante que la moyenne dans les zones où nichent des aigles royaux. A l'inverse, la fréquentation est plus élevée à proximité des habitats prioritaires d'intérêt communautaire, des stations botaniques de chardon bleu et des zones de quiétude pour tétras, démontrant une forte interface entre activités humaines et biodiversité sensible. La fréquentation des zones humides n'est pas significativement différente de la moyenne. Pour les cinq jeux de données testés, certaines zones sensibles ont une fréquentation particulièrement élevée et sont donc à surveiller. Voici par exemple les résultats pour les stations de chardon bleu :



Nom	Fréquentation
Monetier	85.6
Pont de Narreyroux	57.1
Saut du Laire	39.1
Les Clausas	31.5
Lauvitel	27.2
Torrent du Distroit	27.2
Les Rousses	21.3
Le Fournel	11.1
Ruisseau des Berches	8.5
Désert en Valjouffrey	4.9
Combe Méanne	0.4
Ruisseau des Gorges	0.1

Les données actuelles permettent une bonne vision d'ensemble de la fréquentation du PNE, bien que des zones d'ombre demeurent (notamment sur la fréquentation hivernale ou la répartition des différentes activités). Treize recommandations sont présentées dans le rapport pour améliorer le suivi de la fréquentation afin de compléter les connaissances.

Résumé des recommandations

Recommandation n°1 : Harmoniser au maximum les données brutes des jeux de données gérés par le PNE pour qu'ils soient facilement mobilisables. Cela inclus notamment :

- Faciliter le téléchargement centralisé des données
- Inclure dans chaque fichier un unique tableau commençant en A1
- Organiser les tableaux avec une ligne par observation
- Ne pas fusionner de cellules
- Inclure toutes les informations directement dans les tableaux
- Homogénéiser les noms
- Homogénéiser les dates et heures
- Homogénéiser les noms de colonnes
- Différencier les cas sans suivi de fréquentation des cas où la fréquentation est nulle
- Vérifier les données au fur et à mesure de leur acquisition et conserver une version prête à l'emploi

En attendant l'évolution de ces fichiers, le code développé dans le cadre de cette prestation permet d'harmoniser autant que possible les jeux de données et de les rendre interopérables.

Recommandation n°2 : Mettre à jour régulièrement les données externes de fréquentation (grâce aux scripts mis à disposition) et les valoriser au sein du PNE.

Recommandation n°3 : Partager la carte interactive aux personnes intéressées par les données de fréquentation du PNE pour l'utiliser comme vitrine des différents jeux de données.

Recommandation n°4 : Mettre à jour chaque année la carte interactive afin de visualiser les nouvelles données et de partager une version à jour des données disponibles.

Recommandation n°5 : Déployer plus d'éco-compteurs entre 500 et 1000 m et au-delà de 2000 voire 2500 m pour vérifier la corrélation des données à couverture complète sur toute la gamme d'altitude du PNE.

Recommandation n°6 : Obtenir et utiliser les données brutes d'Outdoorvision et Strava serait très intéressant, elles sont une source très riche d'informations.

Recommandation n°7 : A l'avenir, pouvoir cartographier la fréquentation standardisée de différentes saisons (notamment l'hiver) et de différentes activités (en séparant par exemple la randonnée, la course, le cyclisme et le ski) permettrait une étude plus fine de la fréquentation du parc. Cela pourrait se faire par l'analyse des données brutes de Strava et Outdoorvision.

Recommandation n°8 : Déployer des éco-compteurs ou pièges photo dans les bassins versants du Clavens, du Taillefer, du Torrent de la Séveraissette et de la Muande, du Pont-du-Fossé/Ancelle et du Col d'Ornon à Bourg d'Oisans. Le GR54 Col d'Ornon et le Tour du Pays des Ecrins en 6 ou 7 jours peuvent également être intéressants à suivre.

Recommandation n°9 : Mettre en place des suivis hors de la saison estivale (qui concentre la grande majorité des données disponibles) pour pouvoir étudier la fréquentation en intersaison et la fréquentation en période de neige (accéder aux données Strava de ski serait particulièrement intéressant).

Recommandation n°10 : Développer les méthodes de suivi pour des activités de pleine nature autres que la course, la marche et vélo. Pour l'instant aucune donnée n'est prête à l'emploi sur la fréquentation en ski de randonnée, en escalade, en vol libre, ou encore le bivouac.

Recommandation n°11 : Cibler les comportements que le PNE souhaite étudier (par exemple fréquentation des lacs, sorties de sentiers) et développer des méthodes de suivi adéquates, par exemple par piège photo.

Recommandation n°12 : Développer des analyses d'interface entre fréquentation et biodiversité sensible à une échelle fine grâce aux traces GPS issues de Strava ou Outdoorvision.

Recommandation n°13 : Déployer des outils de suivi, notamment par piège photo, dans les sites où la biodiversité est particulièrement sensible à la fréquentation afin d'observer les comportements à fort impact sur la biodiversité (par exemple sortie des chemins ou entrée dans des zones interdites).

Préparation des données

Cette prestation se base sur l'analyse de jeux de données multiples. Certains jeux de données ont été directement récoltés par le PNE ou des partenaires locaux, et sont bien connus du PNE :

- **maisons de parc** : les maisons de parc du PNE enregistrent le nombre de visiteurs par mois, depuis 1986 pour certaines.

- **refuges** : les refuges FFCAM enregistrent le nombre de nuitées commercialisées par mois, principalement en été, depuis 2010.

- **éco-compteurs** : le PNE a mis en place des éco-compteurs, ou compteurs pédestres, sur différents sentiers du parc, pour compter le nombre de passages de piétons. Ils sont principalement actifs en été, les premiers depuis 2006.

- **pièges photo** : le PNE a mis en place des pièges photographiques sur différents sentiers du parc, afin de suivre le passage en caractérisant les activités pratiquées (randonnée, course, vélo...). Ils sont actifs en été et ont été déployés en 2023.

- **enquêtes** : le PNE a mené différentes campagnes d'enquêtes auprès des visiteurs, pour en savoir plus sur leur fréquentation du parc. Les premières enquêtes réalisées en 1980 puis dans les années 2000 (en 2001, 2006 et 2011) n'ont pas été utilisées dans cette prestation car les données brutes n'étaient pas accessibles. Seules les enquêtes de 2022, 2023 et 2024 ont été utilisées.

- **compteurs routiers** : plusieurs routes du PNE ont été équipées de compteurs routiers, qui comptent les véhicules passant chaque heure. Les premiers ont été installés en 1996, et le dispositif s'est arrêté en 2023.

Cette prestation a été l'occasion pour le PNE d'intégrer des jeux de données externes qui peuvent servir à une meilleure quantification de la fréquentation du parc :

- **Outdoorvision** : Outdoorvision est un programme gouvernemental rassemblant les traces GPS d'activités sportives de pleine nature (notamment randonnée, VTT, trail). Le PNE a accès aux données agglomérées en flux sur l'ensemble du territoire du parc.

- **Strava** : Strava est une application privée utilisée pour enregistrer et partager des traces GPS d'activités sportives (notamment la course à pied, le cyclisme ou la randonnée), largement utilisée en France. Le PNE a accès aux données Strava (pour les activités de randonnée et course à pied) agglomérées en flux sur l'ensemble du territoire du parc.

- **iNaturalist** : iNaturalist est une plateforme internationale de saisie de données naturalistes intégrant tous les taxons et ne nécessitant aucune expertise particulière (la saisie se fait par photo ou sons qui pourront être identifiés par des naturalistes experts). De plus en plus répandue en France, plus de 37 000 données ont été recueillies dans le PNE entre 2016 et 2024. La distribution

spatiale de ces données nous informe sur les lieux les plus fréquentés par des randonneurs naturalistes.

- **PlantNet** : PlantNet est une plateforme nationale de saisie de données botaniques gérée par TelaBotanica. Elle est ouverte et utilisée par des amateurs avec différents niveaux d'expertise, grâce à son aide automatisée à l'identification d'après photo. Plus de 8 000 données ont été recueillies dans le PNE depuis 2016. La distribution spatiale de ces données nous informe sur les lieux les plus fréquentés par des randonneurs naturalistes.

D'autres données ont été considérées puis écartées de l'analyse. C'est notamment le cas des données FluxVision, vendues par Orange, qui utilisent les données de téléphonie pour informer sur la fréquentation du parc. Ces données permettent notamment de quantifier les entrées et sorties du parc par période, en fonction de l'origine des visiteurs. Cela peut s'avérer utile pour obtenir des informations à l'échelle du parc, mais ne permet pas de s'intéresser à l'échelle intra-parc puisqu'aucune information spatiale n'est disponible à fine échelle. Or pour l'ensemble des activités présentées dans cette mission, une information spatialisée est nécessaire.

Préparation des données internes

Un élément clé pour la valorisation des données internes de fréquentation est qu'elles soient facilement accessibles et interopérables. Ce n'est pas le cas des données que j'ai reçues du PNE dans le cadre de cette prestation et il m'a fallu plusieurs jours de mise à plat des données pour être prêt à les utiliser. Je détaille ci-dessous la façon dont j'ai traité chaque jeu de données, travail qui est fait sous R et donc répliquable.

A l'avenir, un travail sur le formatage des différents jeux de données dont dispose le PNE pour permettre leur utilisation facile et automatisée serait un grand avantage à toute utilisation des données. Voici quelques grandes règles qu'il convient de respecter pour faciliter l'utilisation des données (illustrée dans la Figure 1) :

- **Faciliter le téléchargement centralisé des données.** Certaines données doivent être téléchargées manuellement pour être obtenues. C'est le cas des fichiers de données éco-compteurs et pièges photo qu'il faut télécharger un après l'autre sur istSOS, en modifiant à chaque fois la date minimum de la période sélectionnée. Cela implique une difficulté d'accès aux données et une possibilité d'erreur (de la part de la personne téléchargeant les données si elle oublie de modifier la date minimum ; ou de la part de la mise à disposition comme je l'ai fait remonté pour les données de pièges photo qui étaient téléchargées avec les mauvais noms de colonnes). La possibilité de télécharger un fichier unique contenant toutes les données des éco-compteurs et un second fichier contenant toutes les données des pièges photo faciliterait l'accès aux données sans erreur.

- **Inclure dans chaque fichier un unique tableau commençant en A1.** Certains fichiers de maisons de parc contiennent deux tableaux sur une même feuille de calcul (voir par exemple la

feuille « Bilan par année » du fichier « GLOBAL FREQUENTATION_MdP_2024.ods »), ce qui complique leur lecture automatisée. Il serait préférable de s'assurer que chaque feuille de calcul ne contient qu'un unique tableau. Il convient également que chaque tableau commence en cellule A1, c'est-à-dire que les noms de colonnes soient inclus dans la première ligne (contrairement aux fichiers de compteurs routiers qui commencent par des lignes de méta-données avant d'inclure le vrai fichier de données).

- **Organiser les tableaux avec une ligne par observation.** Certains fichiers incluent une ligne par observation (par exemple une ligne par jour pour les compteurs routiers) ainsi que des lignes incluant d'autres types d'information (par exemple la moyenne par tranche horaire au cours de l'année). Ces lignes peuvent poser problème si elles sont considérées comme des observations comme les précédentes. Pour éviter ce genre d'erreur, il convient de suivre la logique d'une ligne par observation et de ne pas mélanger les fichiers de données brutes avec les analyses.

- **Ne pas fusionner de cellules.** Certains jeux de données ont des cellules fusionnées, ce qui est absolument illisible dans une lecture automatique des tableaux. C'est par exemple le cas de données de fréquentation de la maison de parc de Briançon pour lesquelles certains mois sont fusionnés (par exemple avril-mai, juin-juillet, et octobre-novembre sont fusionnés en 2018).

- **Inclure toutes les informations directement dans les tableaux.** Certaines informations importantes ne sont actuellement disponibles que dans des métadonnées des fichiers. C'est par exemple le cas des fichiers téléchargés sur istSOS qui n'incluent que la date et le nombre de passages, l'information du nom du compteur devant être ajoutée manuellement. Une bonne pratique est d'inclure systématiquement une colonne avec le nom du compteur, répété pour chaque ligne.

- **Homogénéiser les noms.** Les noms des différents lieux de suivi (noms de compteur, de refuge, de piège photo...) diffèrent souvent entre les fichiers incluant les données de comptage et les fichiers shapefile incluant les coordonnées des sites de suivi. C'est par exemple le cas du refuge de la Bérarde (dénommé « BERARDE » dans le fichier du nombre de nuitées par refuge et « Centre Alpin de la Bérarde » dans le fichier shapefile. Cela demande de faire correspondre les sites manuellement et peut mener à des erreurs. Une bonne pratique est d'harmoniser les noms entre les différents fichiers de données pour que la correspondance soit directe.

- **Homogénéiser les dates et heures.** Les dates peuvent prendre différents formats (par exemple : 20/03/2020, 20/03/20, ou 2020-03-20) et sont parfois mélangées dans un même jeu de données. C'est par exemple le cas dans les comptages routiers où les dates ont des formats différents selon les années. Quand ces formats sont mélangés, cela peut créer des erreurs d'interprétation de dates. Il est donc conseillé de toujours utiliser le même format de date au sein d'un jeu de données, et de faire de même pour les heures.

- **Homogénéiser les noms de colonnes.** Pour les jeux de données divisés entre différents fichiers, il est difficile d'utiliser les fichiers si les noms de colonnes des différents fichiers ne sont pas strictement identiques. C'est par exemple le cas des noms de colonnes des tableaux de compteurs routiers, dont la première colonne varie entre 'Heures et jour' et 'Heures et Jours'

(parfois divisé entre deux lignes), tandis que la dernière colonne varie entre 'TOTAL' et 'Total du jour' (parfois divisé entre deux lignes). Harmoniser les noms des colonnes permet de faciliter la combinaison de ces données. Le fait d'inclure toutes les données dans un tableau unique, comme conseillé plus haut, permet de s'affranchir de ce problème.

- **Différencier les cas sans suivi de fréquentation des cas où la fréquentation est nulle.** Pour certains jeux de données, il est difficile de différencier les cas où le dispositif de suivi n'était pas actif des cas où il était actif mais qu'aucune fréquentation n'a été détectée. C'est par exemple le cas des données d'éco-compteurs, pour lesquelles les heures sans détection de passage sont rapportées avec une valeur 'NaN' (pour « Not a number »), comme on le voit pour la majeure partie des heures de nuit puisqu'il n'y a pas de passage nocturne. Or, on retrouve la même valeur pour des périodes où le dispositif ne fonctionnait pas, comme par exemple l'entièreté du mois de janvier 2019 pour l'éco-compteur d'Adèle Romanche ('NaN' pour la totalité du mois). Il conviendrait d'indiquer une valeur de 0 quand l'éco-compteur fonctionne mais ne détecte aucun passage, et d'indiquer 'NaN' quand l'éco-compteur ne fonctionne pas, afin de faciliter l'utilisation des données.

- **Vérifier les données au fur et à mesure de leur acquisition et conserver une version prête à l'emploi.** Juliette Frigot a passé beaucoup de temps à corriger manuellement les données des éco-compteurs ou pièges photo (qui avaient notamment des problèmes de date) dont certaines avaient été acquises il y a plusieurs années. Ce travail pourrait être fait au fur et à mesure de l'acquisition des données pour que le jeu de données conservé soit propre et prêt à l'emploi. Cela permet notamment de s'assurer que toutes les utilisations de ce jeu de données utilisent les données corrigées.

The diagram shows a table with the following data:

	A	B	C	D
	Nom	Date	Heure	Comptage
1	La Danchère	01/07/25	10:00:00	2
2	La Danchère	01/07/25	11:00:00	10
3	La Danchère	02/07/25	10:00:00	3
4	La Danchère	02/07/25	11:00:00	11
5	Casset	01/07/25	10:00:00	1
6	Casset	01/07/25	11:00:00	0
7	Casset	02/07/25	10:00:00	NaN
8	Casset	02/07/25	11:00:00	15

Annotations:

- Début tableau en A1**: Points to the start of the data table.
- Inclure les noms dans le tableau + correspondance des noms avec le fichier spatial**: Points to the 'Nom' column.
- Unique colonne de comptage**: Points to the 'Comptage' column.
- Aucune cellule fusionnée**: Points to the 'Comptage' column, indicating no merged cells.
- Fichier unique avec tous les compteurs**: Points to the 'Nom' column, indicating a single file for all counters.
- Format date homogène**: Points to the 'Date' column, indicating a consistent date format.
- Pas de ligne "Total" en fin de tableau**: Points to the bottom of the table, indicating no summary row.
- Différencier comptage nul (0) dysfonctionnement (NaN)**: Points to the 'Comptage' column, specifically highlighting the difference between 0 and NaN.

Figure 1 : Exemple de format de données permettant leur prise en main facile. Voir section « Préparation des données internes » pour plus de détails.

Maisons de parc

Les données de visites des maisons de parc sont disponibles à deux résolutions temporelles : comptages annuels depuis 1993, comptages mensuels (pour quelques maisons de parc depuis 2001 et pour la totalité depuis 2012). J'ai utilisé les comptages annuels pour les questions d'évolution temporelle. En revanche, pour comparer les différentes maisons de parc pour une approche spatiale, il faut prendre en compte que la fréquentation de certaines maisons de parc est suivie toute l'année alors qu'elle n'est suivie que l'été pour d'autres. J'ai donc calculé la fréquentation cumulée des mois de juillet et août pour les différentes maisons de parc, entre 2016 et 2024.

Refuges

Les données de nuitées commercialisées dans les refuges ont été fournies dans un fichier centralisé nommé '2024 - ECRINS – PNE.xlsx'. J'ai utilisé la feuille 'evolution année' de ce fichier, à laquelle j'ai manuellement supprimé la première ligne (métadonnées) ainsi que les lignes résumées du bas du tableau. Les noms de refuges ne correspondent pas aux noms du fichier spatial qui m'a été fourni ('refuges.shp'), Juliette Frigot m'a donné les informations pour faire la correspondance entre les deux nomenclatures (voir la nouvelle colonne « Nom_spatial » dans le fichier « Nuitees_refuges_issudocument2024_Juliette.csv »).

Par ailleurs, en 2022, un relevé des destinations des personnes dormant dans 13 refuges du parc a été effectué. Ces données ont été spatialisées par Juliette Frigot (point de départ et d'arrivée des courses) et intégrées à la cartographie interactive.

Eco-compteurs

Les données éco-compteurs ont été téléchargées sur istSOS. Cela implique de cliquer sur chaque compteur un par un, d'ouvrir la page de téléchargement dédié, d'étendre la plage de dates minimum à 2005 (pour être certain de télécharger toutes les données), cliquer sur « Plot » puis sur « Télécharger ». Il faut ensuite ouvrir le fichier et y ajouter une colonne « Nom » dans laquelle on inscrit le nom de l'éco-compteur (qui n'est pas inclus par défaut dans le téléchargement). Cette manipulation manuelle prend du temps et peut mener à des erreurs. Un fichier centralisé avec toutes les données d'éco-compteurs, et incluant une colonne avec le nom de l'éco-compteur, faciliterait grandement l'utilisation de ces données (Figure 1).

Un problème récurrent sur les éco-compteurs est le passage de troupeaux qui peut mener à des comptages journaliers très élevés. Juliette Frigot m'a accompagné dans l'identification des données aberrantes qui pourraient correspondre à des passages de troupeaux et nous avons exclu les journées concernant ces passages. Cela correspond à 5 jours sur l'éco-compteur de Confolens, 3 jours pour La Danchère, 8 jours pour les Gourniers, 1 jour pour le Pré de Mme Carle, 3 jours pour Surette, 6 jours pour Valsenestre, 4 jours pour le Lac de la Douche, 1 jour pour le Lac du Vallon (voir détails dans le script « 0.BTA preparation donnees.R »).

J'ai agrégé ces données par quinzaine en comptant le nombre de passages effectués par demi-mois (première quinzaine entre le 1^{er} et le 15^{ème} jour de chaque mois, seconde quinzaine à partir

du 16^{ème} jour) et en comptant le nombre de jours de fonctionnement du compteur. Pour deux éco-compteurs (Barrière et Serre du Coin), une distinction entre piéton et vélo est possible dans les données. Elle a été utilisée pour la carte interactive.

Pièges photo

Comme les données éco-compteurs, les données de pièges photo ont été téléchargées sur istSOS. En plus de la sélection manuelle de la date, les données pièges photo contiennent différentes variables de comptage (par exemple 'personnes', 'bétail', 'homme', 'femme'...). Pour télécharger ces données, il faut ajouter ces différentes catégories au graphique avant de télécharger les données. Cela prend du temps également et peut mener à des oublis ou des erreurs. De plus, comme je l'ai rapporté à l'équipe, ce système présente actuellement un bug puisque les noms de colonnes ne correspondent pas aux vraies données selon l'ordre dans lequel j'inclus les différentes catégories de téléchargement (on obtient des résultats différents si on demande les colonnes 'personnes' et 'bétail' ou si on demande les colonnes 'bétail' et 'personnes'). Enfin, un traitement manuel de plusieurs fichiers a été nécessaire (réalisé par Juliette Frigot) pour régler certains problèmes de date.

J'ai agrégé ces données par quinzaine en comptant le nombre de passages effectués par demi-mois (première quinzaine entre le 1^{er} et le 15^{ème} jour de chaque mois, seconde quinzaine à partir du 16^{ème} jour) et en comptant le nombre de jours de fonctionnement du compteur.

Enquêtes

Les résultats des quatre campagnes d'enquêtes récentes (2022, 2023, 2024 et Tourond) m'ont été directement envoyés dans quatre fichiers séparés qui sont assez bien harmonisés. La correspondance entre les localisations notées dans les données brutes et le fichier spatial ('lieux_questionnaire.shp') a également demandé un peu de travail manuel. Il a fallu faire la correspondance manuel entre les noms formatés différemment (par exemple 'la_danchere' dans un fichier et 'La Danchere' dans l'autre), ajouter manuellement la localisation de l'enquête faite au refuge Tourond, et rassembler les deux points GPS 'Emparis - Lac lerie' et 'Emparis - Lac noir' puisque dans les fichiers de données des enquêtes, seule la localisation 'emparis' est notée.

Compteurs routiers

Les données de compteurs routiers sont clairement les plus difficiles à prendre en main. Voici les étapes que j'ai suivi pour harmoniser les données :

- Supprimer les fichiers ne contenant pas de données brutes. Certains fichiers contiennent autre chose que des données brutes (fichiers dont le nom inclut 'Répertoire', 'Synthèse', 'Lexique_', 'Table_Correspondance' ou qui sont au format .pdf ou .zip), des données plus spécifiques que celles utilisées ici (fichiers dont le nom inclut '_Clongueurs_' ou 'Supp', fichiers avec données de parking dont le nom commence par 'P '), ou alors des données ajustées manuellement (fichiers dont le nom inclut 'MJM Corrigé'). Tous ces fichiers ont été écartés.

- Supprimer les doublons. Pour certaines années et compteurs, trois fichiers existent : un fichier de comptage dans le sens de la montée, un fichier dans le sens de la descente, et un fichier cumulé. Cette information est souvent indiquée uniquement par la présence d'un 1 dans le nom du fichier (par exemple 'rapportT6_01_June_2018_0050204011-0.xlsx' est le fichier de données dans le sens de la montée, là où 'rapportT6_01_June_2018_0050204012-0.xlsx' est celui dans le sens de la descente et 'rapportT6_01_June_2018_0050204013-0 PrémmeCarle.xlsx' le cumulé). Ne m'intéressant dans cette analyse qu'aux données cumulées, j'ai donc dû supprimer les fichiers détaillant les données de montée ou descente, mais cela uniquement dans le cas où un fichier combiné est disponible, sinon j'ai dû conserver les fichiers de montée et descente et en calculer la somme.

- Lire les tableaux. Les 429 tableaux possèdent une en-tête de plusieurs lignes (variable) qui n'est pas idéale pour lire les données automatiquement. J'ai automatisé l'extraction des données brutes (entre la première colonne de date et la dernière colonne de total, toutes deux possédant des noms variables selon les fichiers ; en excluant les lignes de fin contenant autre chose que des données brutes). Il a également fallu automatiser l'extraction de certaines métadonnées dans l'en-tête du document, notamment le nom du compteur routier, bien que ces informations ne soient pas toujours présentes dans le même endroit du tableau.

- Faire correspondre les noms au fichier spatial. Les noms de compteurs tels qu'ils sont inclus dans les fichiers de comptage varient d'une année à l'autre (par exemple : 'molines-en-champsaur', 'molines -en-champsaur', 'molines cha.', 'molines champ', 'molines en ch' et 'molines en champsaur' désignent le même compteur). De plus, aucun de ces noms ne se retrouve dans le fichier spatial 'Ancien_compteur_routier.shp' qui ne possède que des numéros de compteurs routiers. Il a donc fallu faire une correspondance manuelle, là aussi ralentissant l'utilisation de ces données.

En résumé pour les données de compteurs routiers, il conviendrait de restructurer complètement les données pour que chaque tableau commence dans la cellule A1, que seules les données brutes soient incluses, avec des noms de colonne et des formats identiques d'un tableau à l'autre, et qu'une colonne incluant le nom du compteur (comme il est noté dans le fichier spatial) et le sens de comptage soit ajoutée. La fonction *BTA_PreparerRoutier* créée dans la prestation peut être utilisée comme base pour ce travail.

Recommandation n°1: Harmoniser au maximum les données brutes des jeux de données gérés par le PNE pour qu'ils soient facilement mobilisables. Cela inclut notamment :

- Faciliter le téléchargement centralisé des données
- Inclure dans chaque fichier un unique tableau commençant en A1
- Organiser les tableaux avec une ligne par observation
- Ne pas fusionner de cellules
- Inclure toutes les informations directement dans les tableaux
- Homogénéiser les noms

- Homogénéiser les dates et heures
- Homogénéiser les noms de colonnes
- Différencier les cas sans suivi de fréquentation des cas où la fréquentation est nulle
- Vérifier les données au fur et à mesure de leur acquisition et conserver une version prête à l'emploi

En attendant l'évolution de ces fichiers, le code développé dans le cadre de cette prestation permet d'harmoniser autant que possible les jeux de données et de les rendre interopérables.

Préparation des données externes

Quatre jeux de données externes ont été mobilisés pour l'analyse de la fréquentation : Strava, Outdoorvision, iNaturalist et PlantNet. Ils présentent l'avantage majeur de couvrir l'ensemble du parc, puisqu'ils sont basés sur des données GPS, et peuvent donc nous renseigner sur la fréquentation de l'ensemble des vallées, autant en altitude qu'en fond de vallée.

Les données brutes seront utilisées dans certaines analyses. Pour d'autres, il est nécessaire de transformer ces données en raster d'intensité de fréquentation (carte de chaleur). La méthode pour calculer ces rasters varie légèrement d'un jeu de données à l'autre, mais les rasters calculés ont tous le même format : un raster de résolution de 500x500m couvrant l'ensemble du parc (cœur et zone optimale d'adhésion) avec une projection Lambert 93.

Outdoorvision

Les données Outdoorvision ont été extraites et envoyées directement par l'équipe en charge au Ministère des sports, de la jeunesse et de la vie associative. Les données brutes d'Outdoorvision sont les traces GPS brutes des courses intégrées à la base de données. Ce ne sont pas ces données auxquelles nous avons eu accès (elles sont lourdes et nécessitent un accord avec l'ensemble des fournisseurs de données), mais des données agrégées en flux. L'export qui m'a été transmis intègre les données de randonnée, de course à pied, de cyclisme et de ski de randonnée, sans que ces activités soient différenciées dans l'export.

Les données flux sont des données résumant l'information contenues dans les traces GPS en indiquant le nombre de pratiquants ou de passages réalisés sur un chemin donné. Elles se présentent sous forme d'un fichier vecteur contenant des lignes spatialisées, associées à un nombre de passage, indiquant par exemple que 120 passages ont été répertoriés sur une section de chemin donnée. Les données Outdoorvision ne sont pas agrégées en fonction de chemins existants mais en fonction de la densité des traces (c'est-à-dire qu'un ensemble de traces similaires hors sentier seront agrégées en un chemin bien qu'aucun chemin ne soit connu sur cette zone) contrairement aux données Strava ci-dessous (Figure 2).

Les données flux ne sont pas évidentes à transformer en raster de fréquentation. Cela est dû au fait que deux flux proches peuvent parfois enregistrer des personnes différentes (par exemple quand un chemin se divise en deux chemins parallèles sur une courte distance) ou des mêmes

personnes (par exemple avec deux tronçons consécutifs de chemin). Dans le premier cas, il faudrait calculer la somme des deux tronçons pour estimer la fréquentation, dans le second cas cela reviendrait à compter deux fois les mêmes personnes. Il est ainsi impossible, sans avoir accès aux données brutes (les traces GPS individuelles), de calculer pour chaque cellule du raster le nombre d'utilisateurs présents un jour donné. En revanche, il est possible de calculer une donnée de fréquentation pertinente en multipliant pour chaque tronçon de flux le nombre de passages par la longueur de ce tronçon (en km) et d'ajouter pour chaque cellule les valeurs des tronçons se trouvant dans la cellule. Cette valeur équivaut au nombre de kilomètres parcourus par une personne dans chaque cellule ; elle augmente soit parce que les personnes sont plus nombreuses, soit parce qu'elles effectuent une plus grande distance dans la cellule. Il est à noter que cette approche n'a pas pu être validée par l'équipe d'Outdoorvision qui n'a pas répondu à nos questions et multiples relances au cours de la prestation.

Pour transformer les données flux d'Outdoorvision en raster de fréquentation, j'ai ainsi multiplié pour chaque tronçon le nombre de passages détectés (colonne T des données flux) par la longueur du tronçon en kilomètres. J'ai ensuite calculé pour chaque cellule la somme de ce produit pour l'ensemble des tronçons ayant leur centroïde dans la cellule.



Figure 2 : Exemple de format des données flux de Strava (marron) et Outdoorvision (noir) autour du Refuge du Châtelleret. Les flux Strava sont basés sur des chemins existants, tandis que les flux Outdoorvision sont créés à partir des traces GPS sans comparaison avec les chemins existants. Il en résulte que plusieurs flux Outdoorvision peuvent exister en parallèle là où Strava ne cartographie qu'un chemin (point jaune 1) et des cas où les données Strava sont cartographiées sur l'entièreté du chemin là où Outdoorvision ne le cartographie que partiellement car la fréquentation y est faible (point jaune 2). A l'inverse, Strava n'est pas capable de cartographier des zones fréquentées mais qui n'ont pas de chemin cartographié sur OpenStreetMap, tandis qu'Outdoorvision peut inclure des chemins sur n'importe quelle zone du territoire étudié (point jaune 3), qui peuvent également correspondre à des traces GPS décalées (ce qui ne peut être vérifié qu'en ayant accès aux traces brutes).

Strava

Les données Strava ont été extraites par Juliette Frigot sur le portail Strava auquel le PNE a accès (Strava Metro). Les données brutes sont également des traces GPS brutes des courses réalisées avec l'application. Les données auxquelles nous avons accès sont des données agrégées en flux. L'export qui m'a été transmis intègre les données de randonnée et course à pied. Un téléchargement des données de sport de neige n'était pas possible via Strava Metro.

Contrairement aux données Outdoorvision, l'agrégation des données Strava se fait autour des chemins connus et référencés dans OpenStreetMap (Figure 2). Ainsi, les traces proches des sentiers existants vont y être associées, tandis que les traces qui seraient trop éloignées de tout sentier existant seront perdues (à noter que nous n'avons pas eu accès à la méthodologie détaillée de ces données flux). Les traces brutes peuvent être visibles sur la carte de chaleur du portail en ligne de Strava, mais ne peuvent pas être téléchargées avec l'accès dont le PNE dispose.

Comme pour les données Outdoorvision, il nous est impossible de calculer le nombre de personnes présentes par jour dans une cellule du raster. La même méthodologie a été appliquée ici en multipliant pour chaque tronçon le nombre de passages (colonne « *total_trip_count* ») à la longueur du tronçon (en km) et en sommant ce produit pour l'ensemble des tronçons tombant dans la cellule. Contrairement aux données Outdoorvision, les tronçons Strava sont longs (jusqu'à 1 000m), j'ai donc coupé chaque tronçon aux limites de la grille du raster de 500x500m, afin de ne quantifier que la fréquentation de chaque tronçon de chaque cellule.

Les données Strava peuvent être préparées avec la fonction *BTA_PrepareStrava* développée dans le cadre de cette prestation.

iNaturalist

Les données iNaturalist sont directement téléchargeables sous R grâce à la fonction *get_inat_obs* du package *rinat* (Barve et Hart 2025). Pour accommoder le fait que les données ne sont téléchargeables que par groupe de 10 000 observations au maximum, j'ai coupé le territoire du PNE en une grille de 20 cellules et j'ai téléchargé et regroupé les observations de chacune des cellules. J'ai téléchargé ces données le 19 juin 2025 et je n'ai gardé que les observations des 10 dernières années (depuis 2016).

Les données iNaturalist ne peuvent pas être directement utilisées comme données de fréquentation. En effet, la densité de données iNaturalist est liée à l'intensité de la fréquentation, mais également au nombre d'observations que chaque utilisateur réalise sur un point donné. Pour les utiliser comme données de fréquentation, j'ai créé un raster qui quantifie pour chaque cellule le nombre d'observateurs.jours ayant rapporté des données dans la cellule (une personne rapportant 10 observations dans la même cellule le même jour comptera pour 1 ; une personne rapportant des observations 3 jours différents dans l'année comptera pour 3 ; deux personnes rapportant des observations le même jour comptent pour 2).

Les données iNaturalist pourront être téléchargées et mises à jour grâce à la fonction *BTA_Telecharge_iNaturalist* développée dans le cadre de cette prestation.

PlantNet

Les données PlantNet sont chargées sur la banque mondiale de données de biodiversité GBIF (Global Biodiversity Information Facility) et sont ainsi directement téléchargeables sous R grâce à la fonction *occ_data* du package *rgbif* (Affouard et al. 2025, Chamberlain et al. 2025) en précisant le numéro de jeu de données *7a3679ef-5582-4aaa-81f0-8c2545cafc81*. J'ai téléchargé ces données le 19 juin 2025 et je n'ai gardé que les observations des 10 dernières années (depuis 2016).

Les données PlantNet ont été transformées en densité de fréquentation comme décrit pour les données iNaturalist plus haut (via un raster quantifiant le nombre d'observateurs.jours ayant rapporté des données dans chaque cellule).

Les données PlantNet pourront être téléchargées et mises à jour grâce à la fonction *BTA_Telecharge_PlantNet* développée dans le cadre de cette prestation.

Recommandation n°2: Mettre à jour régulièrement les données externes de fréquentation (grâce aux scripts mis à disposition) et les valoriser au sein du PNE.

Objectif 1 : cartographier la fréquentation à l'échelle du parc

L'objectif principal de cette prestation est d'utiliser l'ensemble des données mobilisables pour étudier la fréquentation du PNE. J'ai d'abord voulu rendre les données accessibles et visualisables grâce à une carte interactive des données de fréquentation. J'ai ensuite réalisé une cartographie standardisée de la fréquentation dans le PNE à deux échelles : un raster de 500x500m et à l'échelle du chemin de randonnée. J'ai également étudié la couverture du PNE par les dispositifs de suivi, afin d'identifier des zones blanches, où les données de fréquentation sont manquantes. Enfin, j'ai exploré la dimension temporelle des données, en utilisant certains jeux de données pour étudier les tendances annuelles et saisonnières de fréquentation.

Cartographie interactive du PNE

J'ai créé une carte interactive permettant de visualiser les différentes données de fréquentation accessibles dans le PNE. Cette carte expose les différents jeux de données internes du PNE (refuges, éco-compteurs, compteurs routiers...) sous forme de marqueurs sur lesquels les utilisateurs peuvent cliquer pour obtenir plus d'informations sur le dispositif de suivi (nom, dates de suivi actif, évolution de la fréquentation...). Cette carte présente également les données externes au PNE telles que les données iNaturalist, ou Outdoorvision, sous forme de données brutes et de raster d'intensité de fréquentation.

Cette carte permet une exploration facile de la diversité des données de fréquentation du PNE et de leur complétude (via la visualisation des mois avec ou sans suivi). Elle est facilement partageable aux différents agents ou partenaires du PNE, et permet ainsi de valoriser au mieux les données du PNE. Enfin, cette visualisation rapide des différents dispositifs de suivi permettra d'identifier certaines erreurs dans les données (des erreurs de localisation, de dates, de valeur de fréquentation...). Elle permet de visualiser :

- les périodes de suivi et les comptages (résumés par quinzaine) pour les éco-compteurs, compteurs routiers, pièges photo en cliquant sur les icônes correspondantes. Plus de détails sont disponibles en passant le curseur sur les graphiques (notamment certains dysfonctionnements ou passages de troupeau pour les éco-compteurs). Pour les éco-compteurs de Barrière et Serre du Coin, deux graphiques sont disponibles pour différencier les piétons et les vélos.
- le nombre de visites annuelles pour les maisons de parc et les refuges.
- le nombre d'enquêtes réalisées par mois.
- la répartition de destinations des courses effectués depuis certains refuges (d'après un relevé effectué en 2022) en ajoutant la couche « Destinations refuges ». L'épaisseur des traits pour les flux représente la fréquence relative de chaque course.
- les données flux (vecteur) et agrégées (raster de 500x500m) de passage Strava et Outdoorvision. L'épaisseur des traits pour les flux représente le nombre de passages de chaque segment.
- les données brutes (point) et agrégées (raster de 500x500m) de densité d'observateurs.jours des données iNaturalist et PlantNet.
- les bassins versants et chemins utilisés dans les analyses de fréquentation standardisée (le nom des chemins et bassins versants est accessible en cliquant dessus).

La carte finale a été fournie en deux versions en rendu de cette prestation. Une première version complète est disponible sous le nom « Carte_interactive_frequentation_PNE.html ». Une seconde version allégée (qui ne contient pas les données flux de Strava et Outdoorvision) est disponible sous le nom « Carte_interactive_frequentation_PNE_sansStravaOutdoor.html » ; cette version est plus réactive et donc plus agréable à utiliser si l'on n'a pas besoin du détail des données flux. Ces deux cartes sont au format html et s'ouvrent donc avec un explorateur internet (par exemple Mozilla Firefox ou Google Chrome) ; cela ne veut pas dire qu'elles sont disponibles en

ligne, les versions fournies étant stockées en local dans le fichier ZIP fourni en rendu de la prestation.

Cette carte a été créée sous R de façon automatisée avec le script « 1a.BTA carte interactive.R » qui intègre la préparation des données, la création de la carte interactive ainsi que la construction et l'enregistrement des deux fichiers html. Tant que le format des données n'évolue pas, elle peut être mise à jour en utilisant ce même script. Le script « 1a.BTA carte interactive.R » n'aura probablement pas besoin d'adaptation pour une mise à jour de la carte, c'est plutôt le script de préparation des données (« 0.BTA preparation donnees.R ») qui devra être mis à jour avec attention, pour être sûr que les nouvelles données soient bien intégrées et que chaque jeu de données soit adéquatement formaté.

Recommandation n°3: Partager la carte interactive aux personnes intéressées par les données de fréquentation du PNE pour l'utiliser comme vitrine des différents jeux de données.

Recommandation n°4: Mettre à jour chaque année la carte interactive afin de visualiser les nouvelles données et de partager une version à jour des données disponibles.

Cartographie standardisée du PNE

L'objectif de cette partie est de créer une carte standardisée de la fréquentation du PNE, qui présente avec le moins de biais possibles les grandes tendances spatiales de répartition de la fréquentation. Pour cela il faut sélectionner les jeux de données pertinents et les combiner de façon à limiter au maximum les biais. Cette carte pourra ensuite être utilisée pour comprendre quels vallons sont les plus fréquentés, les caractéristiques qui sont liées à une forte fréquentation, et comment cette fréquentation interagit avec la présence de zones sensibles pour la biodiversité.

Une méthode initialement considérée puis écartée était d'utiliser des modèles de distribution d'espèces. L'idée générale de ces modèles est d'utiliser des données de présence d'une espèce, d'étudier le lien entre la présence d'une espèce et des covariables (par exemple d'habitat ou d'altitude) et d'utiliser ce lien pour estimer la présence de l'espèce sur l'ensemble du territoire (Miller 2010). Dans le cadre de la fréquentation du PNE, on aurait alors pu modéliser l'intensité de la fréquentation en fonction de l'habitat, de l'altitude, de la distance aux parkings, de la distance aux routes, de la densité de chemins, de la présence de refuge, et des bassins versants. Les données d'entrée auraient été très variées (présence seule avec les données iNaturalist, couverture complète avec les données Strava, échantillonnage avec les éco-compteurs), ce qui pose un défi méthodologique qui peut être traité par l'utilisation de modèles de distribution d'espèces intégrés (Mostert et O'Hara 2023). Ce choix n'a finalement pas été retenu car, contrairement aux données couramment à disposition dans le cadre d'étude de biodiversité, nous avons ici accès à des données couvrant l'intégralité du territoire avec une grande densité de données pour Strava, Outdoorvision, iNaturalist, et PlantNet. Utiliser les modèles de distribution d'espèces, qui se basent sur le lien entre fréquentation et les covariables pour projeter la fréquentation standardisée, mènerait ainsi à une grande perte d'information contenue dans les données brutes. En effet, seule la variation des données brutes qui est expliquée par les covariables choisies serait reportée sur la carte de fréquentation standardisée. Une approche basée sur l'utilisation des données brutes a finalement été préférée pour faire un meilleur usage de la finesse spatiale de ces données.

Méthodes

La méthode retenue pour cartographier la fréquentation standardisée du PNE se base sur les quatre jeux de données à notre disposition couvrant l'ensemble du territoire du PNE : Strava, Outdoorvision, iNaturalist et PlantNet. Ces quatre jeux de données, transformés en rasters de même résolution (500x500m), nous donnent une estimation de la fréquentation de l'ensemble du PNE. Ils peuvent individuellement être considérés comme une représentation de la distribution de la fréquentation. Cependant, ils ne sont probablement pas tous aussi représentatifs de la vraie fréquentation du parc : soit parce qu'ils incluent moins de données et sont donc plus sujets à des phénomènes aléatoires (PlantNet et Outdoorvision sont basés sur moins de données qu'iNaturalist et Strava) soit parce qu'ils incluent des activités plus ou moins représentatives (on peut imaginer que la répartition des données Strava soit biaisée en faveur des secteurs favorables au trail et moins représentative de la distribution des randonneurs ; ou que les données PlantNet se

concentrent sur les randonnées intéressantes d'un point de vue botanique et ainsi peu représentatives des l'ensemble des randonneurs).

Pour comparer ces quatre jeux de données à couverture complète et savoir auxquels faire confiance pour cartographier la fréquentation générale du parc, les éco-compteurs sont très utiles. En effet, les 23 éco-compteurs actifs répartis dans le parc nous donnent accès à une vérité de terrain sur le nombre réel de passages. Cela est notamment permis par le fait que, d'après l'équipe du PNE, les éco-compteurs sont placés de telle sorte que l'on peut considérer que la quasi-totalité des personnes randonnant sur un chemin donné ont été contraintes (par la topographie ou des barrières) à passer sur l'éco-compteur. En comparant les comptages réalisés par les éco-compteurs aux jeux de données à couverture complète, on peut ainsi savoir quels jeux de données sont les plus proches de la réalité des éco-compteurs et favoriser ces jeux de données. Pour ce faire, j'ai calculé la fréquentation estivale des 23 éco-compteurs disponibles en n'utilisant que les mois entre Juin et Septembre après 2020 (les compteurs « Carrelet » et « AdeleRomanche » ont été écartés car ils n'ont pas enregistré de données en 2024). J'ai calculé pour chacun des compteurs restants la fréquentation moyenne en divisant le nombre de passages par le nombre de jours de suivi (Figure 3). Les pièges photo n'ont pas été utilisés dans cette analyse car ils sont tous localisés très proches d'un éco-compteur et portent donc une information similaire ; les utiliser pourrait alors causer de la pseudo-réplication.

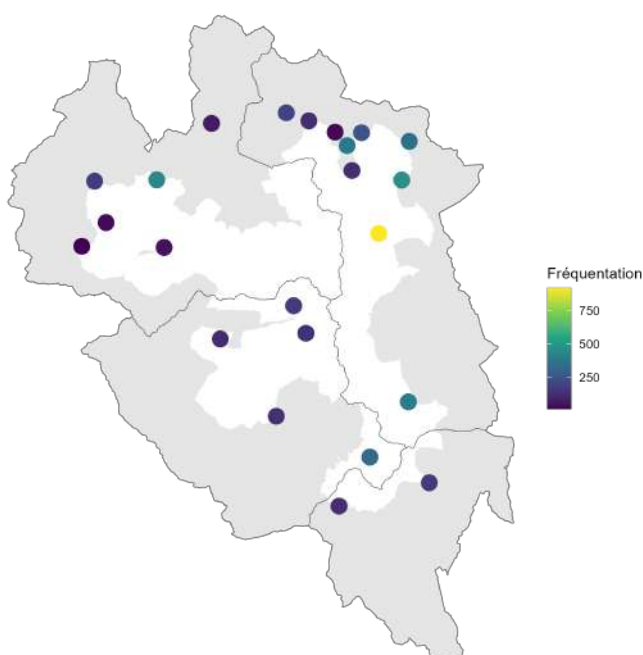


Figure 3 : Fréquentation moyenne estivale des 23 éco-compteurs utilisés pour la cartographie de la fréquentation standardisée. La couleur montre l'intensité de fréquentation estivale mesurée sur chaque éco-compteur comme la moyenne du nombre de passages détectés par jour de suivi. Les lignes noires indiquent les limites des secteurs du parc et la zone blanche indique sa zone cœur.

J'ai ensuite calculé la fréquentation d'après les données Strava, Outdoorvision, iNaturalist et PlantNet sur les chemins sur lesquels sont placés les éco-compteurs. Pour Strava, j'ai isolé le tronçon de chemin sur lequel était placé l'éco-compteur (en prenant le tronçon le moins loin) et j'ai utilisé comme fréquentation de ce tronçon la valeur de la colonne « total_trip_count » pour ce tronçon. Pour les trois autres jeux de données, j'ai calculé la fréquentation autour de l'éco-compteur en utilisant les chemins de Strava. J'ai isolé pour chaque éco-compteur les tronçons qui étaient dans un rayon de 500m autour de l'éco-compteur et qui étaient connectés aux chemins depuis l'éco-compteur (seuil de tolérance de 10m) afin d'écarter les chemins proches mais qui n'étaient pas sur le même itinéraire. J'ai ensuite vérifié manuellement que chacun de ces chemins ne contenait que des chemins passant par l'éco-compteur et j'ai écarté les tronçons qui ne menaient pas nécessairement à l'éco-compteur (par exemple pour la Montée de l'Aigle, j'ai écarté tous les tronçons de l'autre côté de la Romanche puisque les personnes qui s'y trouvent ne vont pas nécessairement passer sur l'éco-compteur de la Montée de l'Aigle). J'ai ensuite calculé la fréquentation Outdoorvision, iNaturalist et PlantNet dans une zone de 100m autour de ces chemins. Pour Outdoorvision, j'ai calculé la somme du produit du nombre de passages par la longueur des tronçons, comme je l'ai calculée pour créer les rasters. Pour iNaturalist et PlantNet j'ai calculé le nombre d'utilisateurs.jours dans cette zone, comme je l'ai calculé pour créer les rasters. A la fin de cette étape, j'ai donc pour chaque éco-compteur une valeur du comptage estival sur l'éco-compteur, une valeur de fréquentation du chemin d'après Strava, une valeur de fréquentation d'après Outdoorvision, une valeur de fréquentation d'après iNaturalist et une valeur de fréquentation d'après PlantNet. J'ai ensuite calculé la corrélation entre les quatre jeux de données à couverture complète et les comptages d'éco-compteurs. Une corrélation parfaite entre les données Strava et les données éco-compteurs, signifierait par exemple que le ratio entre le nombre de traces Strava et les comptages d'éco-compteurs est strictement le même pour chaque éco-compteur (même si le nombre absolu de passages n'est pas le même).

Pour calculer un raster de fréquentation standardisée, j'ai exclu les jeux de données qui auraient une corrélation inférieure à 50 %, considérant qu'ils n'étaient pas suffisamment représentatifs de la fréquentation générale pour être intégrés, et j'ai calculé un coefficient de pondération de chaque jeu de données de valeur $1/(1-CORR)$, où CORR est le coefficient de corrélation. Cette formule permet de pondérer en fonction de l'erreur de chaque jeu de données (c'est à dire sa distance avec une corrélation de 1), donnant ainsi par exemple un poids 4 fois plus important à un jeu de données corrélant à 90 % avec les éco-compteurs (donc avec 10 % d'erreur) qu'à un jeu de données corrélant à 60 % (donc avec 40 % d'erreur). J'ai ensuite normalisé chaque raster corrélé à au moins 50 % (pour qu'ils soient à des échelles comparables), je les ai multipliés par leur coefficient de corrélation respectif, puis je les ai ajoutés. Enfin, j'ai remplacé les 2 % de valeurs les plus hautes par la valeur du quantile à 2 % afin de donner moins d'importance aux rares cellules très fréquentées, et j'ai normalisé l'ensemble des valeurs de fréquentation pour qu'elles s'étalent entre 0 et 100 et puissent ainsi être exprimées en pourcentage.

J'ai également créé une carte de fréquentation à l'échelle des chemins de randonnée principaux du parc annotés par Juliette Frigot (voir section « Cartographie des zones blanches du

suivi interne » pour plus de détails). Contrairement à l'approche par raster qui est basée sur des cellules de même taille, une approche basée sur les chemins doit prendre en compte le fait que les chemins ne font pas la même longueur. Pour Strava et Outdoor, la valeur utilisée pour le raster (nombre de passages multiplié par la longueur des tronçons) n'est plus pertinente puisqu'elle dépend directement de la longueur des chemins. J'ai donc croisé les données Strava et Outdoor avec la cartographie des chemins de randonnées afin de calculer la fréquentation Strava et Outdoor de chaque chemin de randonnée (avec la même méthode que pour le raster) que j'ai ensuite divisé par la longueur du chemin de randonnée. Pour iNaturalist et PlantNet, j'ai calculé le nombre d'observateurs dans un rayon de 100m autour des chemins. Ce nombre d'observateurs augmente nécessairement avec la longueur du chemin puisque la probabilité qu'un observateur fasse une observation augmente avec elle, mais de façon non linéaire. Pour corriger cela, j'ai utilisé les courbes de « species area relationship » qui sont parfaitement indiquées pour ce genre de cas, grâce à la fonction *sar_power* du R package *sars* (Matthews et al. 2019) et j'ai extrait les résidus de ce modèle que j'ai utilisé comme score iNaturalist de chaque chemin. J'ai ensuite, comme pour le raster, normalisé les valeurs des quatre jeux de données de chaque chemin, que j'ai multipliées par leur coefficient de pondération, que j'ai ajoutées, et que j'ai normalisées.

Résultats

Les rasters des quatre jeux de données à couverture complète sont relativement corrélés les uns aux autres, mais avec des différences importantes (Figure 4a). **Strava et Outdoor sont relativement bien corrélés entre eux (73%), tout comme iNaturalist et PlantNet (60%) mais les deux groupes sont moins corrélés entre eux (par exemple 42 % entre Strava et iNaturalist).** Ces corrélations ont été obtenues après avoir rabattu les 2 % de valeurs les plus hautes de chaque jeu de données au quantile à 98 % pour ne pas leur donner trop de poids. Cela indique qu'ils portent bien des informations différentes et qu'il est nécessaire de les comparer et de les combiner.

Les comptages estivaux des éco-compteurs sont corrélés à 84 % avec les données iNaturalist, à 79 % avec les données Strava, à 61 % avec les données Outdoorvision et à 28 % avec les données PlantNet (Figure 5). Ainsi iNaturalist aura un poids de 6,25 dans le calcul de la fréquentation standard, Strava aura un poids de 4,85, et Outdoorvision un poids de 2,54, tandis que PlantNet est écarté car il est corrélé à moins de 50 % aux données des éco-compteurs. Une analyse exploratoire (non présentée ici) n'a pas mis en évidence de structuration spatiale de cette corrélation (ce qui serait le cas si par exemple iNaturalist était plus représentatif dans le Nord du parc et Strava dans le Sud), ce qui justifie que les poids appliqués soient les mêmes sur l'ensemble du territoire. La fréquentation standard résultant de cet exercice suggère une **fréquentation maximale autour des Deux Alpes, du Puy Saint Vincent, du Monêtier les Bains, d'Orcières et d'Embrun** (Figure 4b). **Les secteurs du Pré de Mme Carle, du Gioberney et du Lauvitel sont également particulièrement fréquentés.** La forte densité d'observateurs iNaturalist et PlantNet autour du Lautaret s'explique par la présence de la station de recherche et du jardin botanique qui semblent attirer de nombreux naturalistes. Le même patron apparaît dans la carte basée sur les chemins de randonnée (Figure 6).

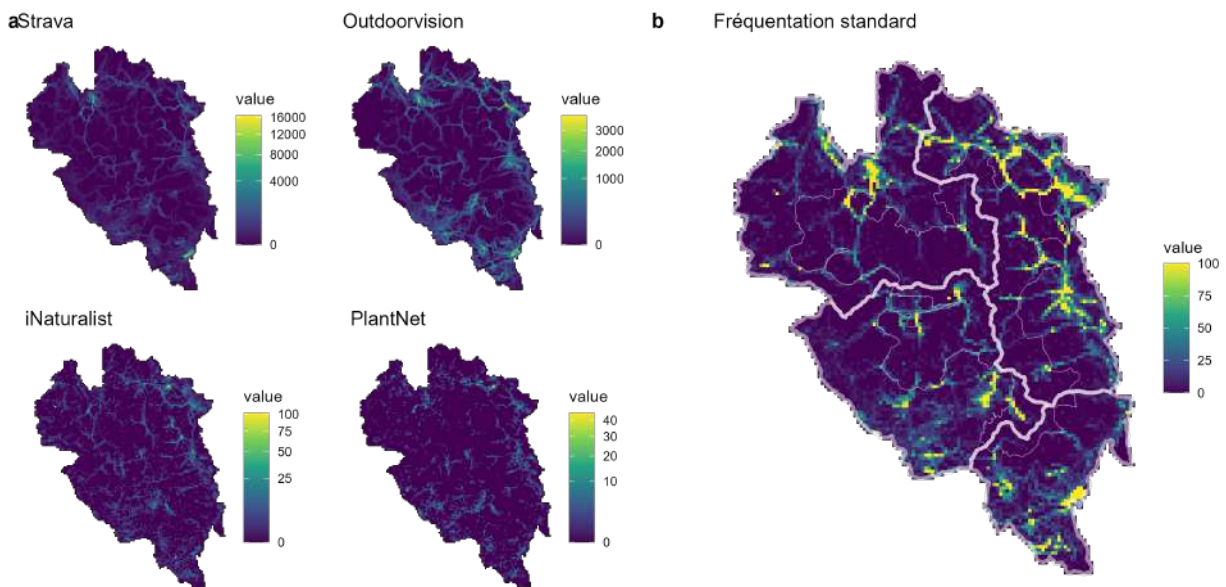


Figure 4 : Données brutes des quatre jeux de données à couverture complète (a) et fréquentation standard calculée pour le Parc national des Ecrins en raster de 500x500m (b). Les lignes blanches sur la carte de droite indiquent les limites des secteurs du parc (ligne épaisse) et de sa zone cœur (ligne fine).

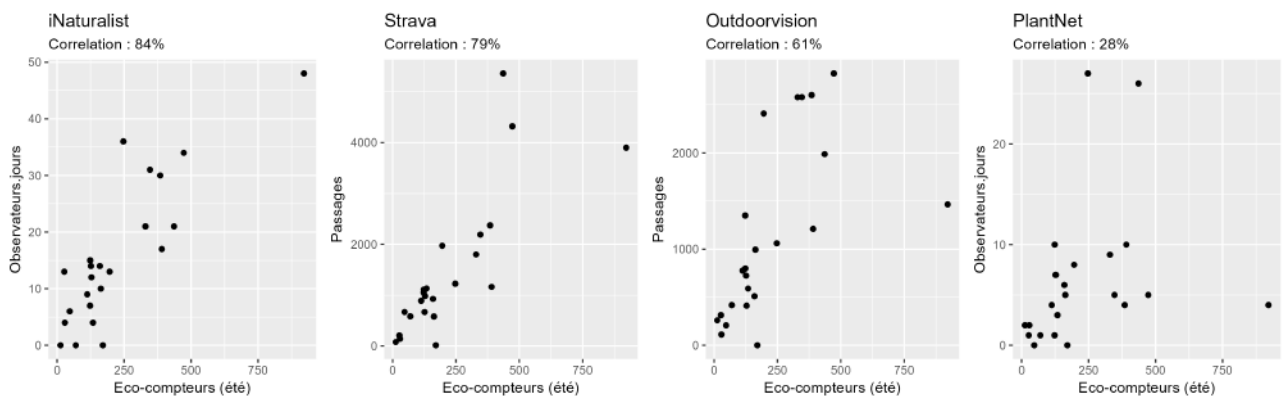


Figure 5 : Corrélation entre les comptages estivaux des 23 éco-compteurs utilisés dans cette analyse et les quatre jeux de données à couverture complète (Strava, Outdoorvision, iNaturalist, PlantNet).

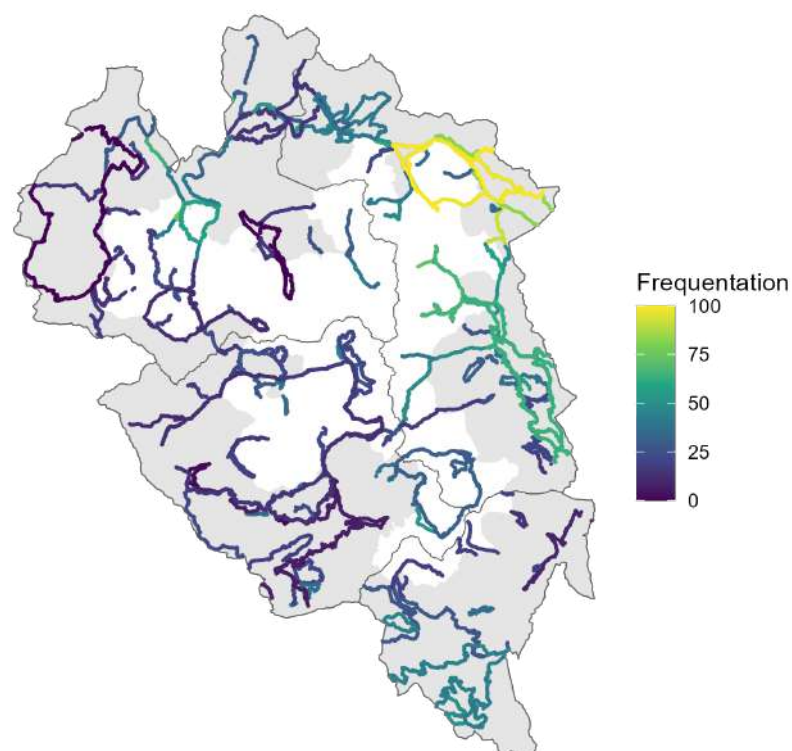


Figure 6 : Fréquentation standardisée des principaux chemins du PNE (basée sur une cartographie manuelle des chemins effectuée par Juliette Frigot). Les lignes noires indiquent les limites des secteurs du parc et la zone blanche indique sa zone cœur.

Une fois cette carte de fréquentation standardisée réalisée, il est possible de l'utiliser pour étudier les facteurs jouant sur la fréquentation du PNE. Par exemple, on peut comparer la fréquentation des différents vallons et voir que **les vallons les plus fréquentés sont ceux du Pont-du-Fossé/Ancelle, du Monétier et de la vallée du Réallon** (Figure 7). On peut également comparer la fréquentation en fonction de variables quantitatives, telles que l'altitude et la distance aux routes, pour s'apercevoir que **les zones les plus fréquentées sont entre 500 et 2500m d'altitude, et proches des routes** (Figure 7).

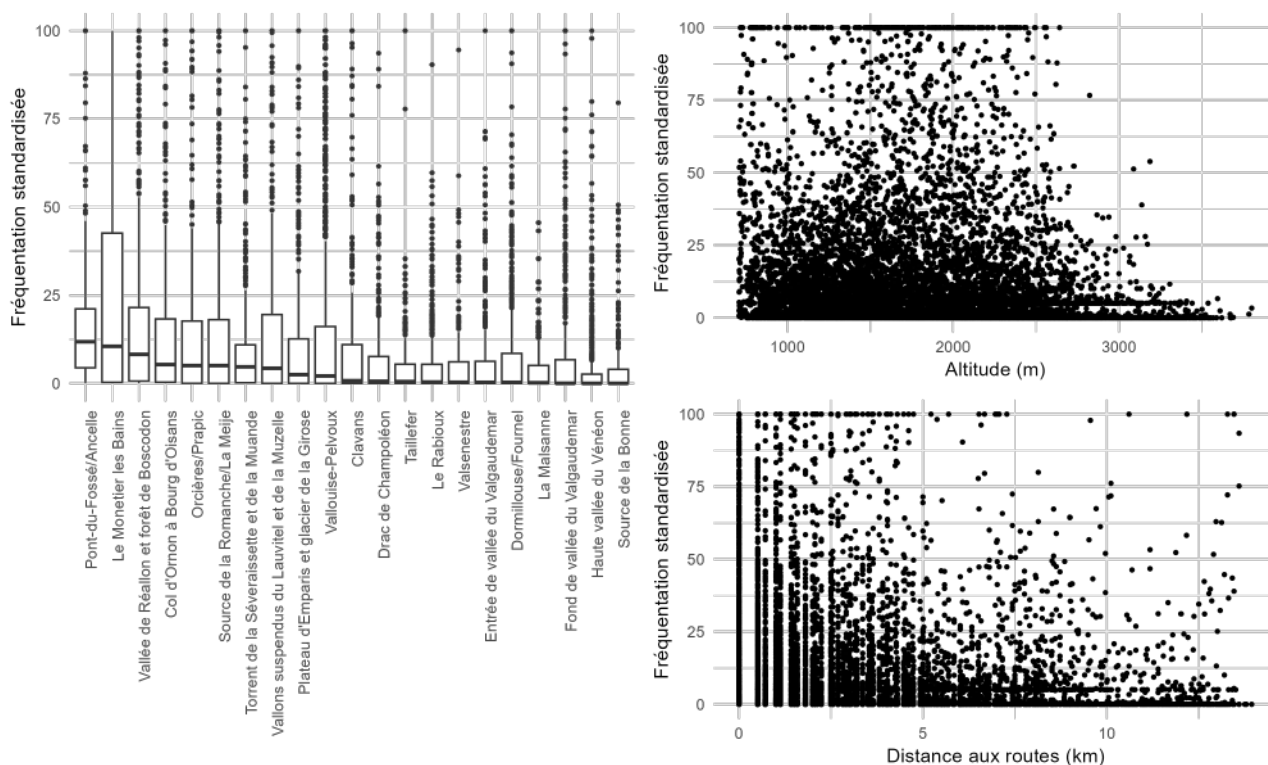


Figure 7 : Corrélation entre la fréquentation standardisée calculée pour le PNE et trois covariables d'intérêt (le bassin versant, l'altitude et la distance aux routes). Pour les boîtes à moustaches (gauche), la ligne horizontale montre la valeur médiane de chaque bassin versant, les limites du rectangle montrent le 1^{er} et 3^{ème} quartile de cette valeur, les lignes montrent la dispersion de ces valeurs et sont complétées par des points pour les bassins versants avec des valeurs extrêmes.

Discussion

La carte de fréquentation standardisée présentée ici intègre trois des quatre jeux de données à couverture complète qui présentent une bonne corrélation avec les données récoltées par les éco-compteurs (iNaturalist, Strava et Outdoorvision). La carte obtenue suggère que les zones les plus fréquentées se trouvent autour des Deux Alpes, du Puy Saint Vincent, du Monêtier les Bains, d'Orcières et d'Embrun. Dans la zone cœur, ce sont les secteurs du Pré de Mme Carle, du Lauvitel et du Gioberney qui semblent les plus fréquentés. Bien que la fréquentation soit particulièrement forte en fond de vallée ou à proximité des routes, on note que certaines zones au-dessus de 2000 m sont tout de même très fréquentées.

La cartographie de fréquentation standardisée présentée ici permet une vision globale de la fréquentation du parc pour comparer les grandes tendances entre vallées ou entre les différentes altitude, à l'échelle de cellules de 500x500m ou à l'échelle des principaux chemins de randonnée. Elle peut également servir à étudier le lien entre la fréquentation du parc et certaines covariables d'intérêt (par exemple la distance aux routes) ou encore à étudier la fréquentation des zones sensibles pour la biodiversité (voir Objectif 3). Il est important de noter que cette carte représente

la fréquentation estivale, puisqu'elle est principalement basée sur des données de course et de randonnée. Réaliser la même carte avec des données hivernales pourrait être très intéressant, mais ces données ne sont pas disponibles pour l'instant (ni pour les éco-compteurs ni pour les données à couverture complète).

Cette cartographie se base sur des hypothèses fortes et présente des limites à considérer. Cette analyse repose fortement sur l'hypothèse que les éco-compteurs nous informent sur la fréquentation réelle des chemins sur lesquels ils sont placés. Le positionnement stratégique des éco-compteurs permet, en principe, d'enregistrer la quasi-totalité des passages sur le chemin en question, mais il est possible que certains défauts des compteurs ou le passage de troupeaux sur certains compteurs n'aient pas été pris en compte. Une autre hypothèse liée aux éco-compteurs est qu'ils sont positionnés de façon représentative dans le parc. Comme étudié dans la partie suivante, leur distribution entre les secteurs est assez bonne bien qu'elle pourrait être améliorée. En revanche, c'est leur distribution altitudinale qui pourrait être améliorée puisque la majorité des éco-compteurs (87 %) sont situés entre 1000 et 2000 m tandis que cette tranche ne représente que 47 % du territoire du parc. A l'inverse, les altitudes inférieures à 1000 m sont sous-représentées dans les éco-compteurs (4 % des éco-compteurs pour 13 % du territoire), tout comme les altitudes supérieures à 2000 m (9 % des éco-compteurs pour 40 % du territoire) puisque l'éco-compteur le plus haut se trouve à 2050 m (éco-compteur des Crevasses). De ce fait, les corrélations que nous avons estimées avec les éco-compteurs ne concernent en réalité qu'une partie du parc, et ne sont peut-être pas valables en fond de vallée ou en haute montagne. Placer de nouveaux éco-compteurs au-dessus de 2000 m voire 2500 m pourrait alors permettre de vérifier que iNaturalist et Strava fournissent de bonnes données de fréquentation sur toute la gamme d'altitude du parc, et pas seulement dans la zone des 1000-2000 m. Par ailleurs, il serait intéressant à l'avenir de cartographier l'incertitude de cette carte de fréquentation, mais aucune méthode n'a été identifiée pour l'instant pour calculer de façon fiable l'incertitude (qui doit intégrer l'incertitude liée à chaque jeu de données, ainsi que l'incertitude liée aux poids des différents jeux de données).

Enfin, il faut noter que cette carte est basée sur des données mélangeant plusieurs activités (randonnée, trail, footing en bas de vallée, cyclisme) que nous ne pouvons pas différencier. La méthode appliquée ici donne plus de poids aux activités détectées par les éco-compteurs, donc typiquement le trail et la randonnée. Idéalement à l'avenir, il serait bien de pouvoir cartographier les différentes activités séparément. Obtenir les données brutes de Strava et Outdoorvision (c'est à dire les traces GPS plutôt que les flux agrégés) pourrait permettre de mieux séparer les différentes activités, et permettrait par la même occasion de créer un raster basé sur le nombre de personnes.jours passant dans chaque cellule, ce qui le rendrait plus facilement comparable avec les autres données.

Recommandation n°5: Déployer plus d'éco-compteurs entre 500 et 1000 m et au-delà de 2000 voire 2500 m pour vérifier la corrélation des données à couverture complète sur toute la gamme d'altitude du PNE.

Recommandation n°6: Obtenir et utiliser les données brutes d'Outdoorvision et Strava serait très intéressant, elles sont une source très riche d'informations.

Recommandation n°7: A l'avenir, pouvoir cartographier la fréquentation standardisée de différentes saisons (notamment l'hiver) et de différentes activités (en séparant par exemple la randonnée, la course, le cyclisme et le ski) permettrait une étude plus fine de la fréquentation du parc. Cela pourrait se faire par l'analyse des données brutes de Strava et Outdoorvision.

Cartographie des zones blanches du suivi interne

Bien que les données externes dont dispose le PNE (Outdoorvision, iNaturalist, etc) couvrent de manière relativement homogène l'ensemble du PNE, les données internes sont elles issues d'un échantillonnage non homogène. La couverture des données d'enquête, d'éco-compteurs, de compteurs routiers, de nuitées en refuge, de visite des maisons du Parc, et de pièges photos n'étant pas homogène, il est probable que certaines zones soient mieux suivies que d'autres. Dans cette partie, l'objectif est de détecter les zones blanches, c'est-à-dire des zones qui manquent de données directes de fréquentation, alors qu'elles sont relativement fréquentées.

Méthodes

Dans cette analyse, nous cherchons à identifier les vallons ou chemins qui ne sont pas suivis directement par le PNE et dans lesquels il serait pertinent d'ajouter des dispositifs de suivi. Il nous a semblé qu'une carte basée sur des unités spatiales (bassins versants et chemins) était plus pertinente pour cette analyse qu'une carte basée sur une grille homogène (type raster) pour identifier les zones blanches.



Figure 8 : Découpage du PNE en 21 bassins versants utilisés pour la cartographie des zones blanches.

J'ai donc basé cette analyse sur deux cartes : une à résolution large basée sur les bassins versants et une à résolution plus fine basée sur les chemins de randonnée. Pour les bassins versants, j'ai utilisé la carte nationale des bassins versants topographiques (IGN et OFB 2025), j'ai écarté les bassins versants périphériques dont moins de 50 km² était inclus dans le PNE et j'ai divisé le bassin versant du Vénéon en deux en suivant la limite entre les communes St Christophe en Oisans et les Deux-Alpes. La carte résultante contient 21 bassins versants (Figure 8). Pour compléter cette analyse à une échelle plus fine, j'ai croisé la carte des chemins de plus de 1000m du PNE avec la carte des bassins versants (N=272 chemins) que Juliette Frigot a renommé et rassemblé en 165 chemins pertinents pour le PNE (Figure 9).

Pour chacune de ces unités spatiales j'ai compté le nombre de dispositifs de suivi locaux disponibles pour le PNE parmi les éco-compteurs, les pièges photos, les compteurs routiers, les refuges comptant les nuitées, et les maisons de parc (Figure 10). Pour l'analyse par bassin versant j'ai compté le nombre de dispositifs de suivi dans chaque bassin versant, tandis que pour l'analyse par chemin j'ai compté le nombre de dispositifs de suivi dans une zone tampon de 50m autour du chemin. Seuls les dispositifs toujours en fonctionnement en 2024 ont été utilisés (par exemple les compteurs routiers arrêtés en 2020 ont été exclus) puisque nous souhaitons identifier les zones où le suivi de la fréquentation actuelle est insuffisant. Les données d'enquêtes n'ont pas été considérées ici puisqu'elles ne nous offrent pas d'informations quantitatives (le nombre de personnes interrogées n'est pas nécessairement corrélé à la fréquentation) et qu'elles ne sont pas réalisées chaque année.

Cette simple somme permet d'identifier les zones qui sont peu suivies. Un manque de suivi n'est pas réellement problématique s'il s'agit d'une zone très peu fréquentée. A l'inverse une zone qui est fortement fréquentée mais dans laquelle le PNE n'a aucun suivi fin de la fréquentation peut constituer une priorité pour l'installation de nouveaux dispositifs de suivi (par exemple d'éco-compteurs ou pièges photo). Pour identifier cela, j'ai utilisé la carte de fréquentation standard calculée dans la partie précédente. J'ai ensuite comparé pour chaque unité spatiale (bassins versants ou chemins) la fréquentation moyenne au nombre de dispositifs de suivi de la fréquentation.

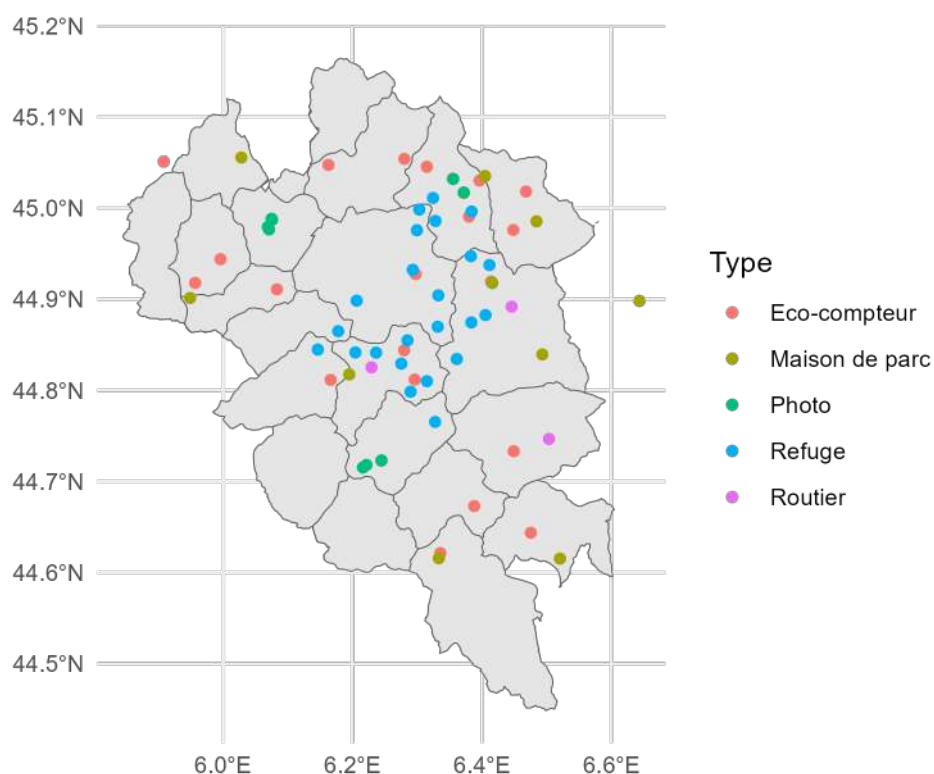


Figure 10 : Carte des dispositifs de suivi interne au PNE par bassin versant.

Résultats

Les dispositifs de suivi mis en place dans le PNE sont principalement **concentrés dans la zone cœur du parc, notamment dans le Fond de vallée du Valgaudemar, Vallouise-Pelvoux, et la Source de la Romanche/La Meije** (Figure 11a ; Tableau 1). A l'inverse, **10 bassins versants ne sont suivis que par un ou deux dispositifs, et quatre bassins versants ne comptent aucun dispositif : les vallons de Clavens, du Taillefer, du Torrent de la Séveraissette et de la Muande et du Pont-du-Fossé/Ancelle**. Hormis le Taillefer qui est peu fréquenté, les trois autres bassins sans dispositifs sont relativement fréquentés, tout comme le Col d'Ornon à Bourg d'Oisans et Orcières/Prapic alors qu'ils n'ont qu'un dispositif de suivi (Figure 11b, Figure 11c). Ces bassins versants peuvent être identifiés comme des priorités pour le déploiement de nouveaux dispositifs de suivi.

Puisque les éco-compteurs sont si importants dans l'élaboration de la carte de fréquentation standardisée, on peut également s'intéresser à la répartition spécifique de ces compteurs. **Aux quatre bassins versants qui ne comptaient aucun dispositif de suivi s'ajoutent les bassins versants du Col d'Ornon à Bourg d'Oisans et de la Source de la Bonne qui ne contiennent aucun éco-compteur** (Figure 12 ; Tableau 1). Parmi les bassins les plus fréquentés, trois vallons n'ont qu'un éco-compteur et pourraient donc être de bons candidats pour en accueillir un second : les vallons suspendus du Lauvitel et de la Muzelle, la Vallée de Réallon et forêt de Boscodon, et Vallouise/Pelvoux. A l'inverse, le bassin versant des Sources de la Romanche / La Meije compte cinq éco-compteurs.

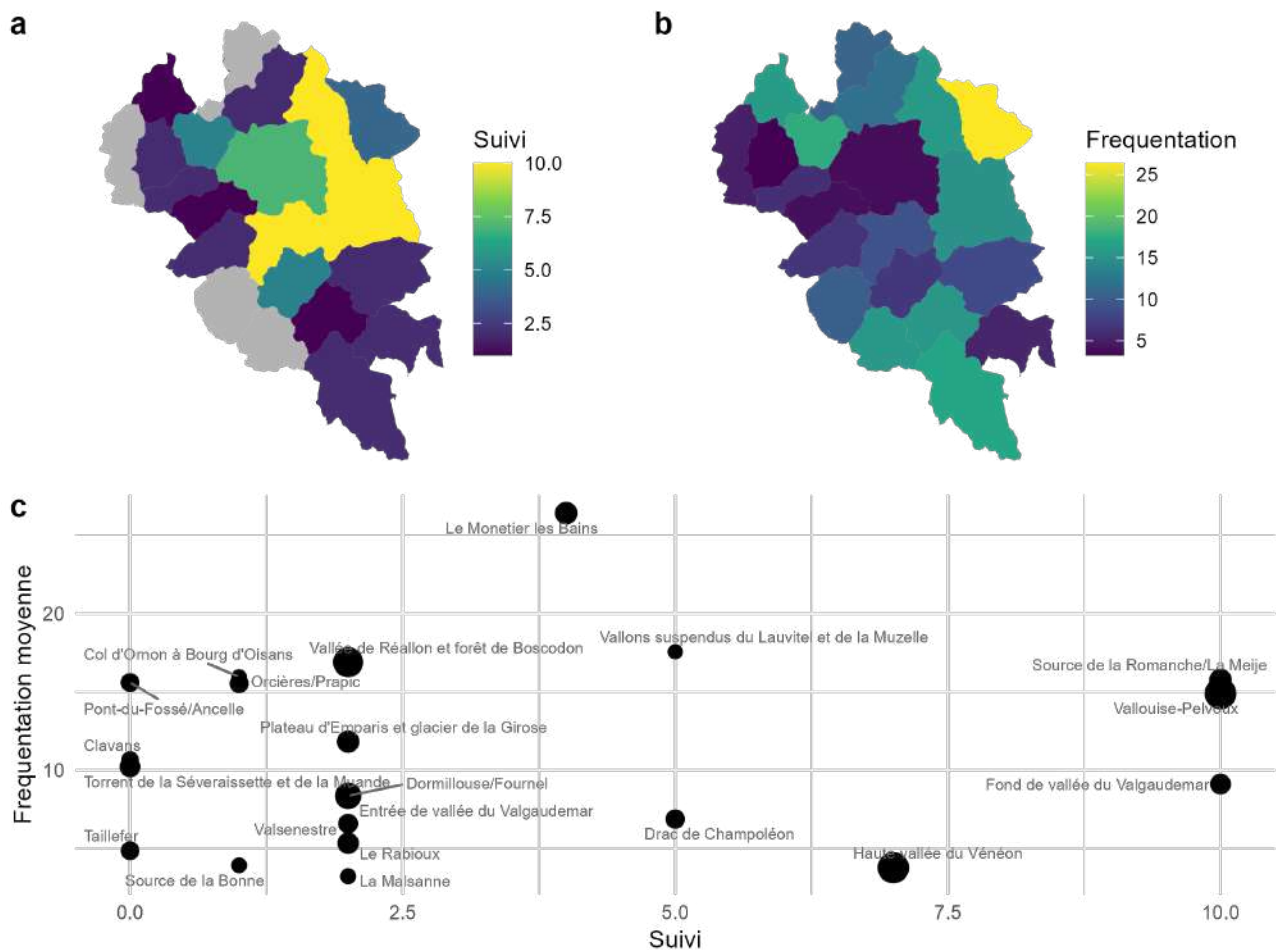


Figure 11 : Cartographie des zones blanches à l'échelle du bassin versant. a : intensité de suivi (nombre de dispositifs de suivi par bassin versant ; les bassins en gris n'ont aucun dispositif de suivi) ; b : intensité de la fréquentation à l'échelle du bassin versant ; c : fréquentation en fonction de l'intensité du suivi par bassin versant. Les noms des bassins versants sont disponibles dans la Figure 8 et sur la carte interactive.

Tableau 1 : Nombre de dispositifs de suivi (Eco-compteurs, compteurs routiers, pièges photo, refuges, maisons de parc ; somme dans la colonne « Suivi ») et fréquentation moyenne des différents bassins versants.

Bassin versant	Suivi	Eco-compteurs	Compteurs routier	Pièges photo	Refuges	Maisons de parc	Fréquentation
Source de la Romanche/La							
Meije	10	5	0	2	3		0 15.8
Vallouise-Pelvoux	10	1	1	0	6		2 14.9
Fond de vallée du Valgaudemar	10	2	1	0	6		1 9.1
Haute vallée du Vénéon	7	1	0	0	6		0 3.8
Vallons suspendus du Lauvitel et de la Muzelle	5	1	0	4	0		0 17.6
Drac de Champoléon	5	1	0	3	1		0 6.9
Le Monétier les Bains	4	2	0	0	0		2 26.4
Vallée de Réallon et forêt de Boscodon	2	1	0	0	0		1 16.9
Plateau d'Emparis et glacier de la Girose	2	2	0	0	0		0 11.8
Dormillouse/Fournel	2	1	1	0	0		0 8.4
Entrée de vallée du							
Valgaudemar	2	1	0	0	1		0 6.6
Valsenestre	2	1	0	0	0		1 6.1
Le Rabioux	2	1	0	0	0		1 5.3
La Malsanne	2	2	0	0	0		0 3.2
Col d'Ornon à Bourg d'Oisans	1	0	0	0	0		1 16.0
Orcières/Prapic	1	1	0	0	0		0 15.5
Source de la Bonne	1	0	0	0	1		0 3.9
Pont-du-Fossé/Ancelle	0	0	0	0	0		0 15.6
Clavans	0	0	0	0	0		0 10.7
Torrent de la Séveraissette et de la Muande	0	0	0	0	0		0 10.2
Taillefer	0	0	0	0	0		0 4.9

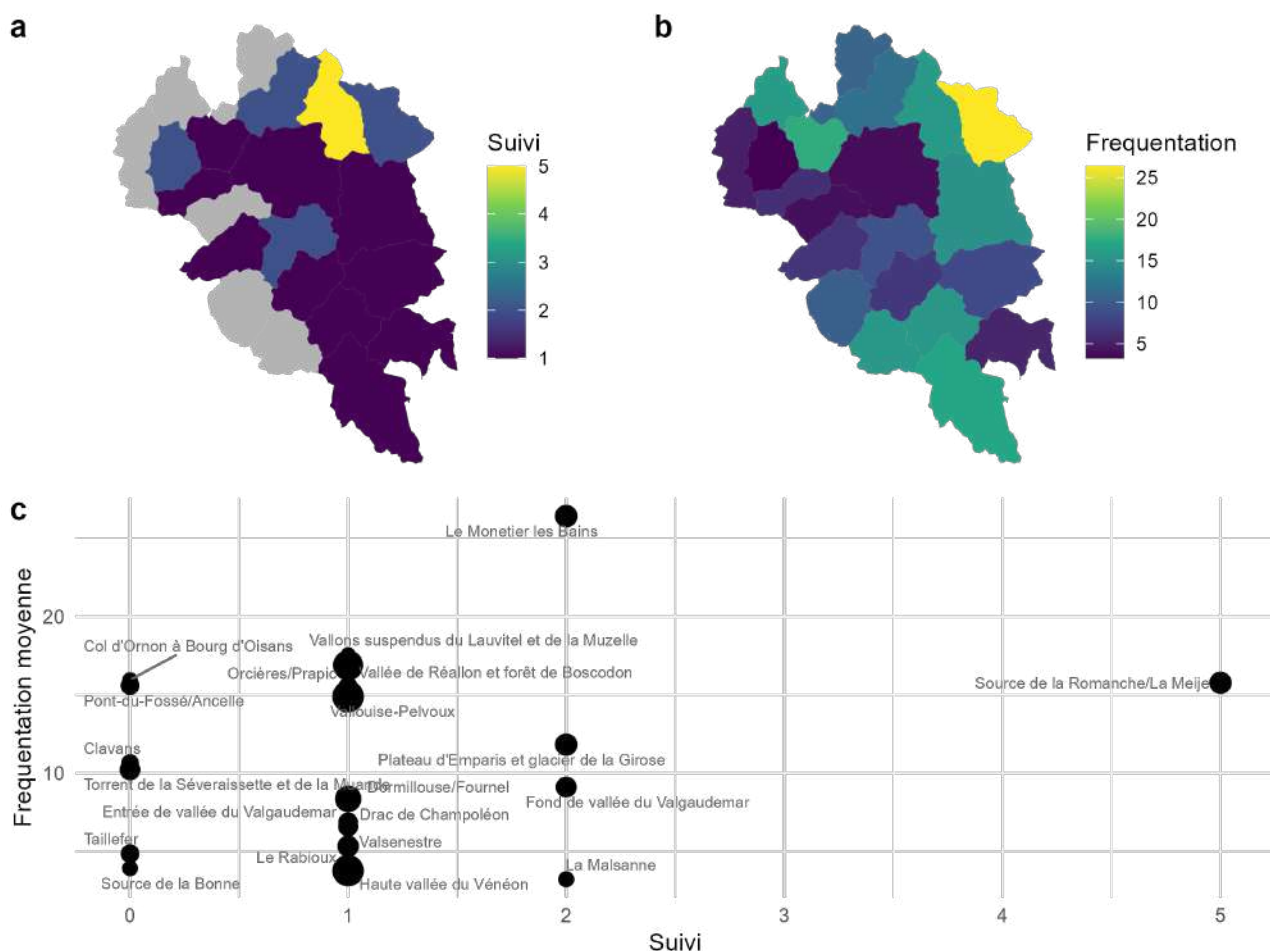


Figure 12 : Cartographie des zones blanches à l'échelle du bassin versant. a : nombre d'éco-compteurs (les bassins en gris n'ont aucun éco-compteur) ; b : intensité de la fréquentation à l'échelle du bassin versant ; c : fréquentation en fonction de l'intensité du suivi par bassin versant. Les noms des bassins versants sont disponibles dans la Figure 8 et sur la carte interactive.

L'échelle des bassins versants permet d'identifier les gros trous dans les dispositifs de suivi mais il est utile de la compléter par une analyse plus fine pour vérifier que l'ensemble des grands chemins de randonnée sont suivis dans le PNE. **Les chemins de randonnée les moins suivis dans le PNE sont ceux en zone d'adhésion** (Figure 13a), notamment au Sud et à l'Est du parc. En particulier, **le GR54 - Col d'Ornon, le Tour du Pays des Ecrins** (en 6 ou 7 jours) et la **Cascade de la Pissette** sont très fréquentés (au delà de 60 %) mais n'ont aucun dispositif de suivi déployé sur leur parcours (Figure 13c).

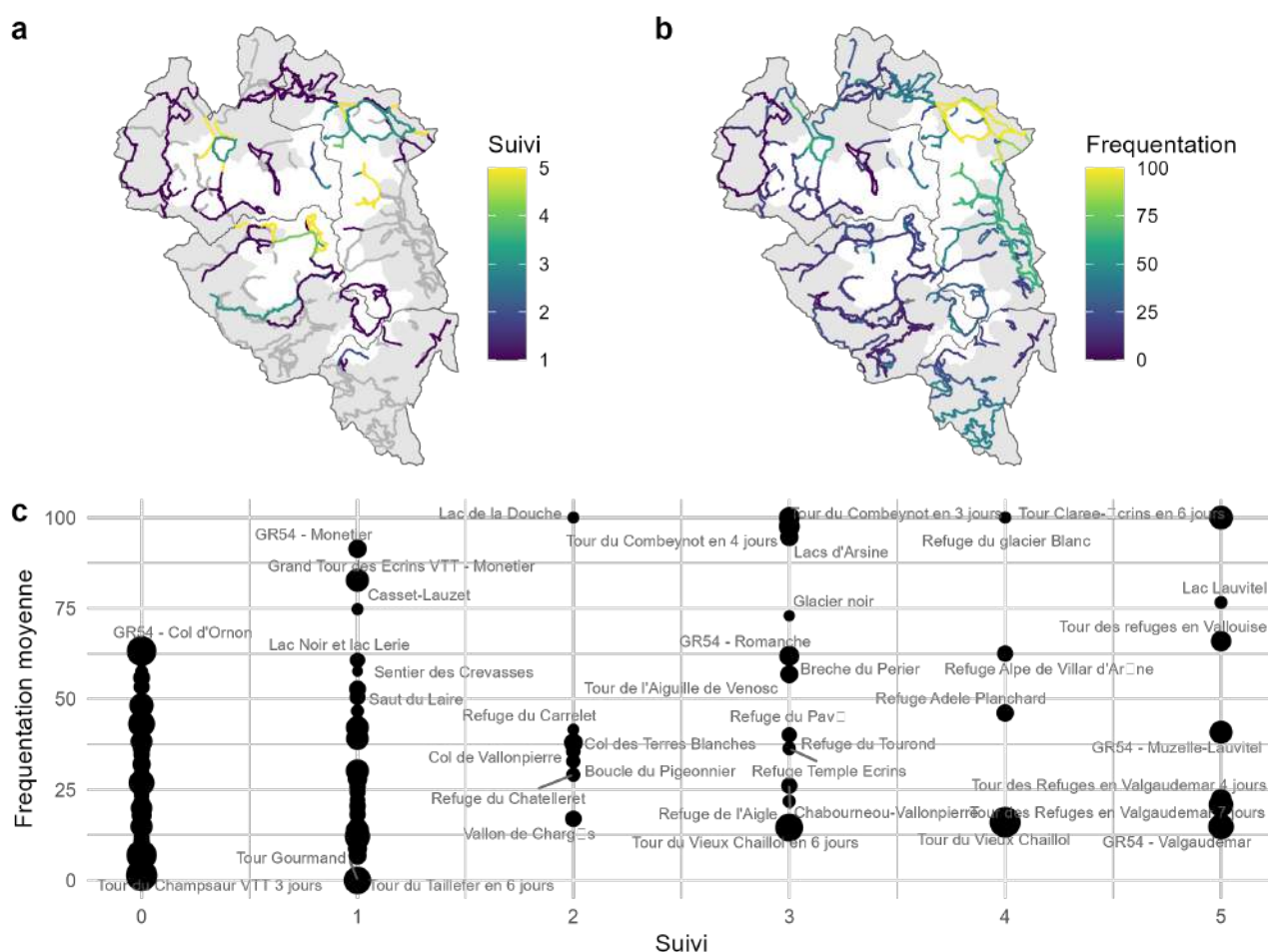


Figure 13 : Cartographie des zones blanches à l'échelle du chemin de randonnée. a : intensité de suivi (nombre de dispositifs de suivi par chemin; les chemins en gris n'ont aucun dispositif de suivi) ; b : intensité de la fréquentation à l'échelle du chemin ; c : fréquentation en fonction de l'intensité du suivi par chemin. Les noms des chemins sont disponibles dans la Figure 9 et sur la carte interactive. Le nombre de dispositifs de suivi par chemin est donné en Annexe 1.

Discussion

L'analyse des zones blanches des dispositifs de suivi permet de mettre en évidence quatre bassins versants (sur un total de 21) qui n'incluent aucun dispositif de suivi, et deux bassins versants supplémentaires qui ne contiennent aucun éco-compteur. Ces six bassins versants peuvent être considérés comme des priorités pour déployer de nouveaux dispositifs de suivi. Une fois que chaque bassin versant sera couvert par au moins un éco-compteur, il sera alors possible de chercher à combler les zones blanches à une échelle plus fine, celle des chemins de randonnée. L'analyse permet d'identifier les chemins les plus fréquentés sur lesquels aucun dispositif de suivi n'est déployé, comme le GR54 Col d'Ornon et le Tour du Pays des Ecrins en 6 ou 7 jours.

Les conséquences de ce manque de suivi sont diverses. D'abord, cela indique que le parc, par ses données internes, a des zones d'ombre sur la fréquentation de bassins versants entiers de son territoire. De plus, dans le cadre des analyses quantitatives présentées ici, la représentativité

territoriale des dispositifs de suivi est très importante, notamment celle des éco-compteurs dans le cadre de la cartographie de la fréquentation standardisée. Pour ce rôle là, déployer des éco-compteurs sur deux ou trois années peut déjà s'avérer intéressant et suffisant pour vérifier la corrélation entre les passages réels et les données externes. Enfin, les dispositifs de suivi déployés par le parc sont les seuls données utilisables pour une approche temporelle (voir partie « Analyse temporelle de la fréquentation ») et donc pour étudier l'évolution de la fréquentation du parc au fil des ans. Il est donc important que ces données puissent couvrir l'ensemble du parc, afin d'estimer des tendances représentatives de l'ensemble de la fréquentation du PNE.

Si le PNE investit dans le déploiement de nouveaux dispositifs de suivi dans le parc, les bassins versants et chemins identifiés ici peuvent être considérés comme candidats prioritaires. Les dispositifs particulièrement intéressants à ajouter seraient des éco-compteurs ou des pièges photo qui peuvent être utilisés pour une approche quantitative de la fréquentation de randonnée. A terme, il serait également intéressant de mieux étudier la complémentarité des différents dispositifs de suivi pour mieux valoriser la diversité des données récoltées par le parc.

Recommandation n°8: Déployer des éco-compteurs ou pièges photo dans les bassins versants du Clavens, du Taillefer, du Torrent de la Séveraissette et de la Muande, du Pont-du-Fossé/Ancelle et du Col d'Ornon à Bourg d'Oisans. Le GR54 Col d'Ornon et le Tour du Pays des Ecrins en 6 ou 7 jours peuvent également être intéressants à suivre.

Analyse temporelle de la fréquentation

Plusieurs suivis réalisés par le PNE permettent d'étudier la tendance inter-annuelle de la fréquentation. De plus, certains jeux de données externes peuvent être utilisés pour étudier les variations saisonnières de la fréquentation. Dans cette section je présente des exemples d'analyses possibles pour explorer ces variations temporelles de la fréquentation.

Tendance inter-annuelle

L'analyse temporelle la plus classique consiste à étudier l'évolution d'une donnée au cours du temps en cherchant à mettre en évidence une évolution linéaire (augmentation ou diminution). Plusieurs jeux de données internes permettent d'étudier l'évolution de la fréquentation dans le PNE : les éco-compteurs, les compteurs routiers, les nuitées en refuge, et les visites de maisons de parc. Les jeux de données externes, en revanche, ne le permettent pas puisqu'ils sont issues d'applications récentes qui ont vu leur popularité augmenter dans les dernières années ; les tendances seront donc nécessairement à la hausse, pas parce que la fréquentation a augmenté mais parce que le recours à ces applications a augmenté.

Pour les quatre jeux de données utilisés, j'ai appliqué un modèle linéaire mixte sur les données annuelles (refuges et maisons de parc) ou agrégées par quinzaines (éco-compteurs et compteurs routiers ; en ne gardant que les quinzaines entre Juin et Septembre avec au moins 10 jours de fonctionnement de l'appareil) avec l'année comme variable réponse principale. J'ai utilisé la fonction *glmer.nb* du R package *lme4* (Bates et al. 2015), supposant une distribution en binomiale négative des données (idéale pour les données de comptage un peu trop dispersées) et en appliquant un effet aléatoire au nom du refuge, de la maison de parc ou du compteur. Pour les éco-compteurs et les compteurs routiers, j'ai ajouté dans le modèle un effet sur le mois (permettant d'expliquer la différence de fréquentation entre les mois de juin, juillet, août et septembre) et sur le nombre de jours de fonctionnement du compteur durant la quinzaine.

Les résultats indiquent une **diminution de la fréquentation pour les données de visites de maisons de parc, les compteurs routiers, et les éco-compteurs** (Figure 14). A l'inverse, les nuitées en refuge semblent augmenter, mais cela peut être dû à l'augmentation de leur capacité plutôt qu'à une vraie augmentation de la fréquentation. De plus, on observe que les mois de **Juillet et Août sont bien plus fréquentés que les mois de Juin et Septembre** (Figure 14).

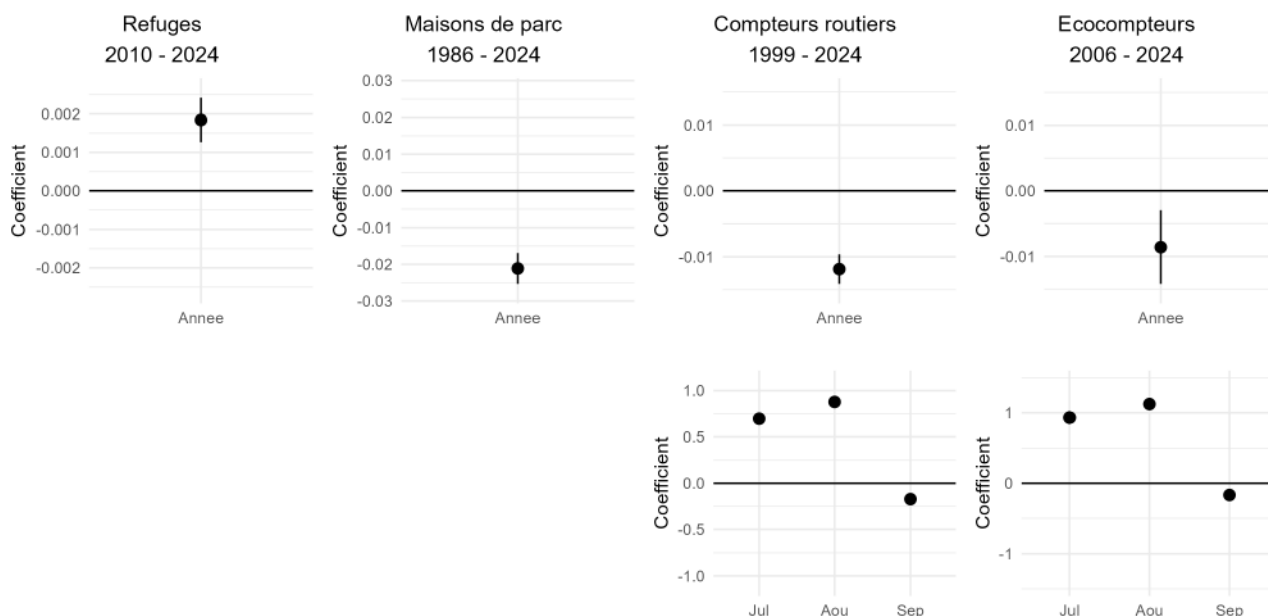


Figure 14 : Estimation des 4 modèles de tendance annuelle pour les nuitées en refuge, les visites de maisons de parcs, les compteurs routiers, et les éco-compteurs. Les graphiques du haut montrent l'effet de l'année sur la fréquentation (négatif si la fréquentation décline) et les graphiques du bas montrent l'effet des mois pour les compteurs routiers et les éco-compteurs (le mois de Juin est utilisé comme référence ; un coefficient positif indique donc une fréquentation plus importante qu'au mois de Juin). Le nombre d'années considérées dans l'analyse temporelle est indiqué dans le sous-titre de chaque colonne. Les barres d'incertitude correspondent à l'erreur standard du coefficient multipliée par 1,96.

Il est éventuellement possible d'ajouter une analyse pour chercher des points d'inflexion dans les tendances positives grâce à des modèles spécifiques, comme ceux proposés par la fonction *chnngptm* du R package *chnngpt* (Fong et al. 2017). Il suffit pour cela d'appliquer ces modèles sur les résidus du modèle annuel. De tels seuils pourraient par exemple être suspectés autour de 2020 si les pratiques de montagne avaient évolué drastiquement en réponse à la période de confinement due au Covid-19. De tels seuils ne sont toutefois pas visibles dans les données brutes et, par conséquent, ne ressortent pas comme significatifs des modèles explorés.

Analyse saisonnière

Un autre type d'analyse permet d'étudier la saisonnalité de la fréquentation pour connaître les pics et les périodes creuses de fréquentation. Il faut pour cela des données couvrant l'ensemble des saisons, ce qui n'est pas le cas de la majorité des données utilisées dans cette analyse. J'ai appliqué cette analyse sur les données iNaturalist.

J'ai utilisé toutes les données iNaturalist depuis janvier 2017, en calculant sur le parc le nombre d'observateurs pour chaque jour. J'ai ensuite utilisé la fonction *decompose* du R package *stats* (R Core Team 2025) qui permet de décomposer les observations d'une série temporelle en trois composantes : la tendance générale, l'effet saisonnier (identique d'une année à l'autre), et les variations aléatoires (Figure 15). On voit que les données iNaturalist sont de plus en plus nombreuses puisque la ligne « trend » montre une croissance continue. L'effet saisonnier semble également très marqué.

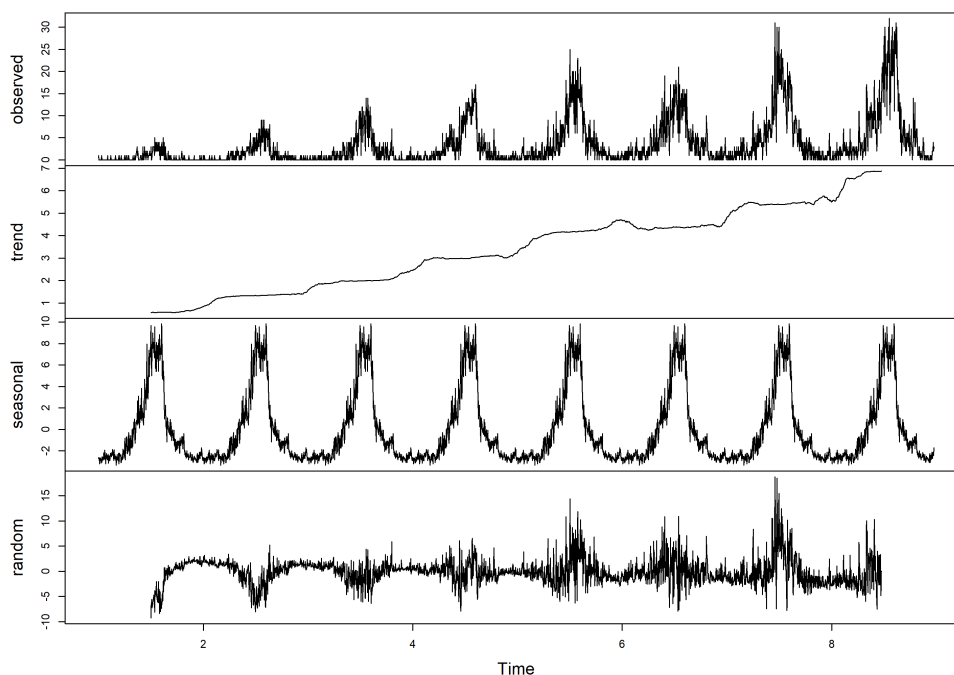


Figure 15 : Décomposition de la série temporelle des données de iNaturalist (nombre d'observateurs par jour sur le territoire du PNE) entre tendance (« trend »), effet saisonnier (« seasonal ») et variations aléatoires (« random »).

Il est ensuite possible d'extraire l'effet saisonnier (effet identique d'une année sur l'autre) pour l'étudier plus précisément. J'ai ainsi extrait le nombre d'observateurs journaliers estimé par la décomposition et j'ai découpé trois périodes (Figure 16). Le pic de fréquentation correspond à la période englobant tous les jours au-dessus du quantile à 95 % de fréquentation (c'est-à-dire que cette période inclura les 5 % des journées les plus fréquentées) ; ce **pic de fréquentation s'étale du 18 Juin au 9 Août**. A l'inverse, j'ai identifié la période de très basse fréquentation, comme la période englobant les 50 % de jours les moins fréquentés ; cette **période de basse fréquentation s'étale du 22 Septembre au 30 Avril**. Entre les deux se trouve la période de fréquentation moyenne.

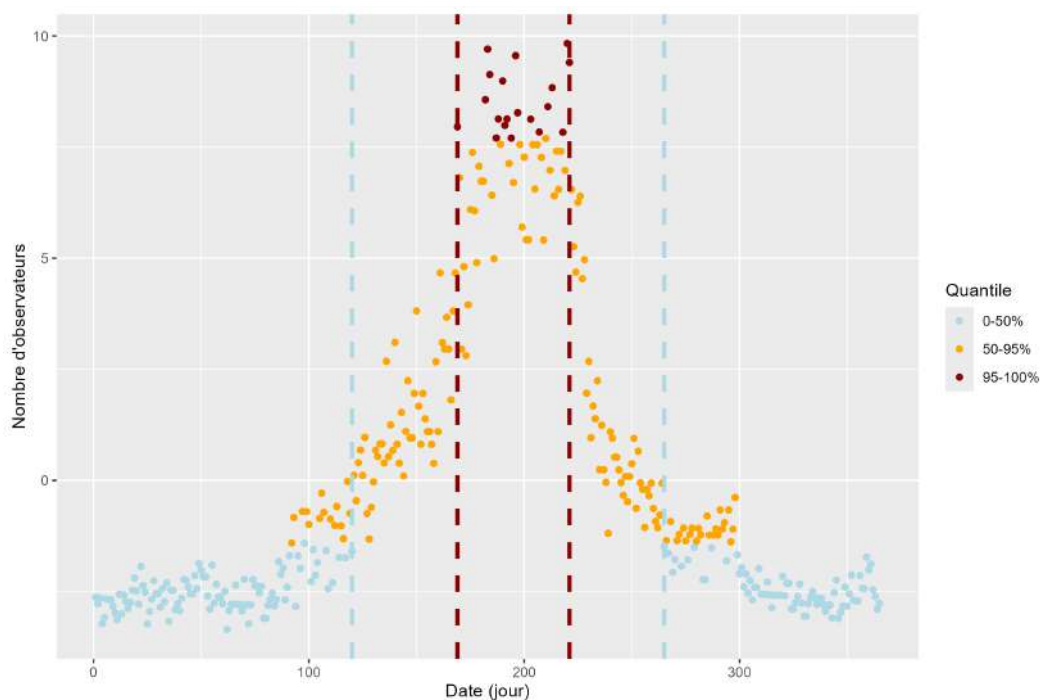


Figure 16 : Effet saisonnier du nombre d'observateurs iNaturalist dans le PNE. La couleur des points indique leur quantile (par exemple les jours 5 % de jours avec le plus d'observateurs sont colorés en rouge). Ces quantiles ont été utilisés pour découper en trois périodes de fréquentation : pic de fréquentation (entre les deux barres rouges), fréquentation moyenne (entre les barres bleues et rouges), et basse fréquentation (à l'extérieur des barres bleues) ; voir texte pour plus de détails.

Discussion

Ces analyses montrent que les données de fréquentation interne au parc sont bonnes pour étudier les tendances de la fréquentation sur les dernières décennies. Il faut dans ce cadre là prêter une attention particulière à ce que représente le jeu de données. Par exemple, une baisse de la fréquentation des maisons de parc ne traduit pas nécessairement une baisse générale de la fréquentation, mais peut-être un changement d'habitude des visiteurs. De la même façon, une hausse de fréquentation des refuges peut être induite par une augmentation de leur capacité d'accueil. Les données de compteurs routiers et d'éco-compteurs, toutes deux montrant une baisse de la fréquentation, sont probablement des témoins plus fiables de l'évolution de la fréquentation.

Les données disponibles pour une analyse de la saisonnalité de la fréquentation sont plus rares puisque la majorité ne sont collectées qu'en été. Les données iNaturalist sont de qualité suffisante pour étudier la saisonnalité, mais présentent un biais important qui est la saisonnalité des organismes observés par les naturalistes. En effet, à fréquentation égale, une journée d'été verra plus d'observations iNaturalist renseignées puisque le nombre de plantes ou animaux disponibles pour l'observation est bien plus important. Il est donc possible qu'une partie de la saisonnalité observée dans la Figure 16 soit due à une plus grande difficulté à faire des observations naturalistes. Cela dit, cet effet est probablement atténué par le fait que nous nous

intéressons ici au nombre d'observateurs et pas au nombre d'observations. Il suffit donc qu'un observateur note au moins une observation dans le parc un jour donné pour être compté, ce qui arrive fréquemment même en hiver. Cela illustre encore une fois la difficulté d'étudier la fréquentation hivernale du PNE.

Recommandation n°9 : Mettre en place des suivis hors de la saison estivale (qui concentre la grande majorité des données disponibles) pour pouvoir étudier la fréquentation en intersaison et la fréquentation en période de neige (accéder aux données Strava de ski serait particulièrement intéressant).

Objectif 2 : étudier la fréquentation spécifique du vallon du Lauvitel

Le PNE dispose de plusieurs données mobilisables à une échelle plus fine, sur un site spécifique d'intérêt. Il est en effet possible d'utiliser les données collectées par le parc telles que les données issues des pièges photos, mais également d'utiliser les données externes telles que Strava ou iNaturalist à une échelle plus fine. Le second objectif de la prestation est d'explorer quelles questions peuvent être étudiées avec les données actuelles et de réfléchir aux données supplémentaires à collecter pour étudier d'autres questions.

Au vu de la disponibilité des données et de l'intérêt de ce vallon en termes de protection, nous avons choisi de nous concentrer sur le vallon du Lauvitel. L'objet n'est pas ici de mener une analyse poussée et complète de la fréquentation de ce vallon, mais plutôt d'utiliser ce vallon comme exemple de ce qui pourrait être étudié pour de multiples sites du PNE. Les limites géographiques de ce vallon ont été fournies par le PNE.

Au sein de ce vallon, nous disposons de plusieurs jeux de données (Figure 17) :

- les données Strava, à une échelle temporelle fine (heure par heure)
- les données Outdoorvision
- les données iNaturalist (N=543) et PlantNet (N=113)
- les données de 4 pièges photos positionnés dans le vallon
- les données d'un éco-compteur et d'un compteur routier situés à l'entrée du vallon
- les données d'enquête collectées à l'entrée du vallon (N=726)

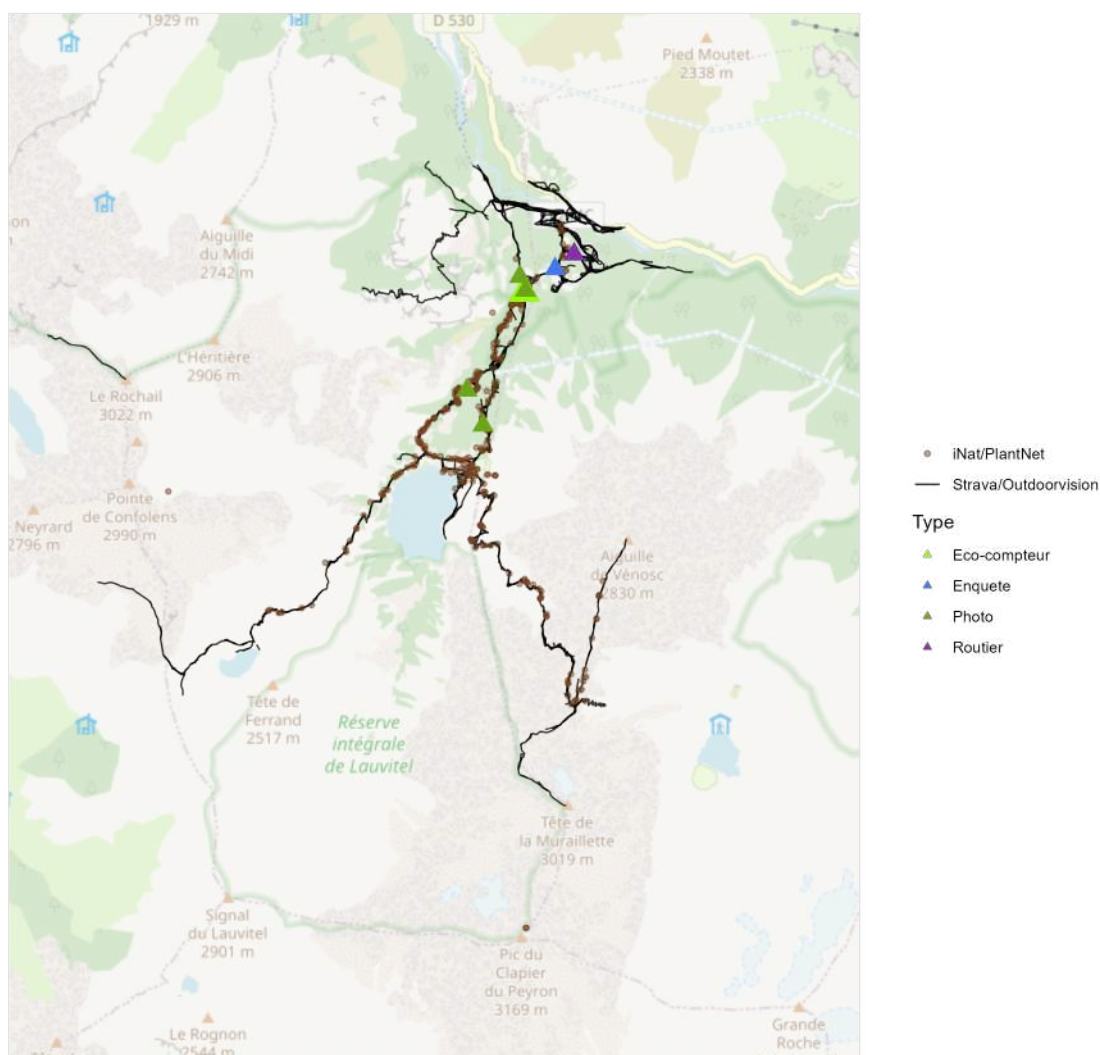


Figure 17 : Données disponibles pour le vallon du Lauvitel. Seules les données iNaturalist et PlantNet avec une incertitude de localisation de moins de 20m ont été rapportées sur la carte.

Diffusion spatiale dans le vallon

A l'échelle du vallon, il peut être intéressant de connaître la diffusion spatiale de la fréquentation : comprendre comment les randonneurs se dispersent dans le réseau de chemin du vallon et comment cela se traduit en terme d'altitude.

Les données Strava sont parfaitement indiquées pour étudier cette diffusion fine. En effet **les données brutes de comptages de passages de Strava sont directement utilisables pour observer la diffusion de la fréquentation** (Figure 18a). On y voit par exemple que le chemin montant au lac par l'Est est plus fréquenté que celui montant par l'Ouest, qu'une bonne partie de la fréquentation s'arrête au lac du Lauvitel, et que pour ceux qui continuent la randonnée plus loin que le lac, le col du Vallon est une destination plus fréquentée que la montée vers la Brèche du Périer et le lac de Plan Vianney.

Pour quantifier cette diffusion, j'ai restreint la carte de diffusion à la partie au-dessus du parking de la Danchère. Ainsi cette carte n'inclut que la montée vers le lac du Lauvitel et au-delà,

avec deux points de départ : le parking (segment entre la Danchère et le ruisseau du Lauvitel) et les Gauchoirs (dernier segment du chemin longeant le ruisseau pour arriver à la Danchère). J'ai calculé la somme des passages de ces deux points de départ (que j'utilise comme un nombre total de randonnées effectuées dans le vallon) et j'ai divisé le nombre de passages de chaque segment par cette somme. Cette valeur est un pourcentage de fréquentation pour chaque segment, par rapport au nombre de randonnées effectuées dans le vallon (Figure 18b). Les valeurs agrégées par Strava ne permettent pas une transformation parfaite en pourcentage, ce qui se traduit par le fait que des chemins opposés peuvent avoir une somme de plus de 100 % (par exemple au début du chemin on a 71 % rive droite et 34 % rive gauche ce qui fait une somme de 104%) ; il faut donc garder en tête que ces valeurs sont approximatives.

On observe que **92 % des randonneurs du Lauvitel utilisant Strava arrivent de la Danchère tandis que 8 % arrivent des Gauchoirs** (Figure 18b). **La proportion de personnes commençant cette randonnée qui vont jusqu'au lac est d'environ 100 %, environ 60 % par le chemin à l'Est et 40 % par le chemin à l'Ouest.** La proportion des personnes continuant au-delà du lac est en revanche bien inférieure, avec **environ 5 % qui continuent vers la Brèche du Périer et le lac de Plan Vianney et environ 20 % qui continuent vers le col du Vallon.** Le **col du Vallon semble particulièrement emprunté pour un passage depuis ou vers la vallée de la Muzelle** puisque la descente vers le lac de la Muzelle semble presque autant empruntée que la descente vers le lac du Lauvitel (17 % contre 20 %).

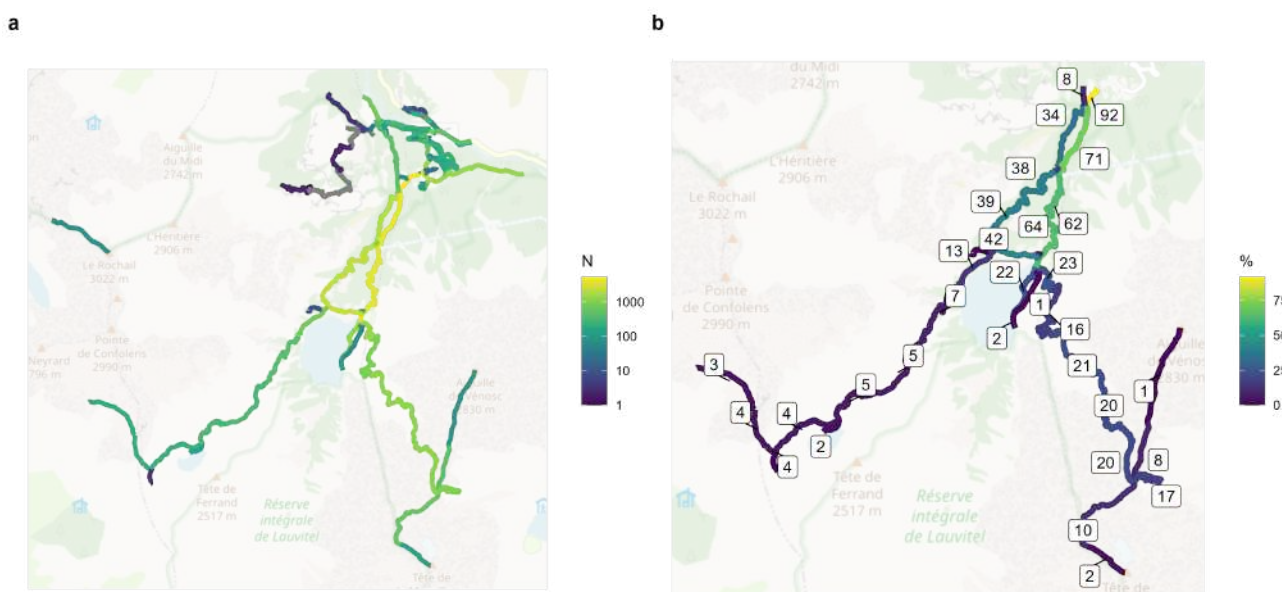


Figure 18 : Diffusion des données Strava dans le Lauvitel, tous chemins confondus (a) ou en se focalisant sur les chemins au-dessus du parking (b). Dans le graphique b, la fréquentation est ramenée en pourcentage approximatif de fréquentation des départs de randonnée (arrivée du parking + arrivée par les Gauchoirs).

Un autre type de diffusion qui peut être étudié dans le vallon du Lauvitel est la diffusion autour des chemins, pour comprendre si les randonneurs restent sur les sentiers ou s'en écartent. Pour cela, on peut utiliser les données iNaturalist. Une incertitude de localisation est souvent renseignée avec les données iNaturalist (colonne « public_positional_accuracy ») et les données PlantNet (colonne « coordinateUncertaintyInMeters ») ; je n'ai gardé pour cette analyse que les 252 données iNaturalist et 42 données PlantNet précises à moins de 20 m. J'ai ensuite calculé la distance de chaque point au chemin le plus proche (en utilisant les données Strava et en écartant les trois points localisés sur des chemins non tracés dans les données Strava). On observe que **la grande majorité des points (88 %) sont localisés dans une zone de 20 m autour du chemin**, qui correspond à la limite d'incertitude de localisation utilisée ici, on peut donc considérer que ces points sont localisés sur le chemin. Les **12 % restants sont localisés entre 20 et 50 m du chemin (6%), entre 50 et 100m (4%) ou entre 100 et 150 m (2%;** Figure 19). Les points hors sentier sont **principalement localisés à proximité du lac de Lauvitel, à des altitudes d'environ 1500 m**. Une limite importante de cette approche est que certains observateurs iNaturalist donnent parfois les coordonnées de l'animal qu'ils observent plutôt que leur propre localisation ; dans ces cas-là une observation considérée ici comme hors sentier peut avoir été réalisée depuis un chemin.

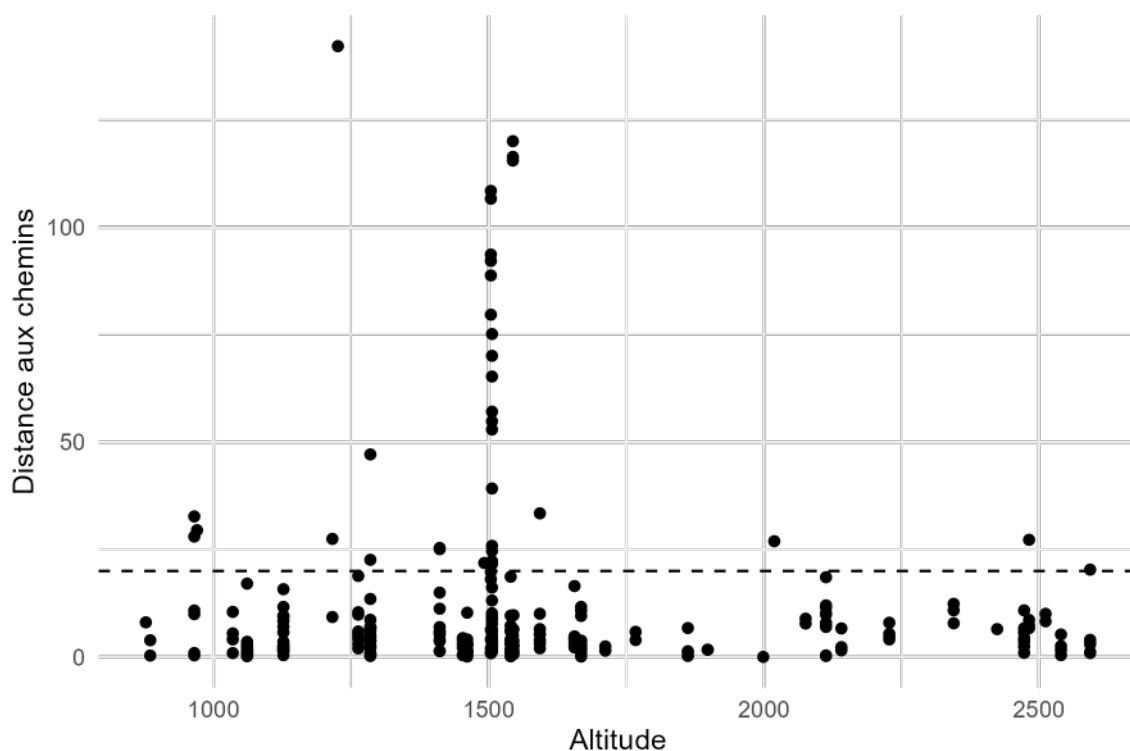


Figure 19 : Distance aux chemins des observations iNaturalist précises (incertitude < 20 m) du Lauvitel en fonction de leur altitude.

Distribution temporelle fine

Plusieurs jeux de données sont disponibles avec une grande finesse temporelle, puisque les données d'éco-compteurs et pièges photographiques, de compteurs routiers, les données Strava et iNaturalist sont disponibles heure par heure. Cela permet d'étudier la distribution temporelle du vallon afin de déterminer les périodes où la fréquentation est la plus importante. Cette distribution temporelle fine peut par exemple nous renseigner sur les heures de moindre dérangement de la biodiversité, même s'il faut noter que les activités de bivouac et la présence en refuge n'est pas représentée par ces données (puisque les traces GPS s'arrêtent en général quand les randonneurs arrivent au point d'arrivée) donc que des dérangements peuvent perdurer au-delà de ces horaires.

Pour cela, j'ai utilisé une autre extraction des données Strava, réalisée spécialement pour le vallon du Lauvitel, qui inclut les données heure par heure. J'ai restreint aux mois d'été (Juin à Septembre) et j'ai calculé pour chaque tronçon le nombre de passages par heure ainsi que l'altitude de chaque tronçon. Cela permet d'étudier la distribution temporelle des passages en fonction de l'altitude (Figure 20).

La fréquentation du vallon s'étale principalement de 7h à 17h. Les hautes altitudes (2000m et plus) sont très vite atteintes avec un pic de fréquentation à partir de 8 ou 9h, et voient leur fréquentation baisser plus tôt dans l'après-midi. Dans les altitudes plus basses, on voit une première vague de fréquentation assez tôt, qui correspond aux personnes rejoignant les hautes altitudes, et une seconde vague (entre 10h et 16h) qui semble plutôt correspondre à un public faisant une randonnée moins longue et qui s'arrête à 1500m au lac du Lauvitel.

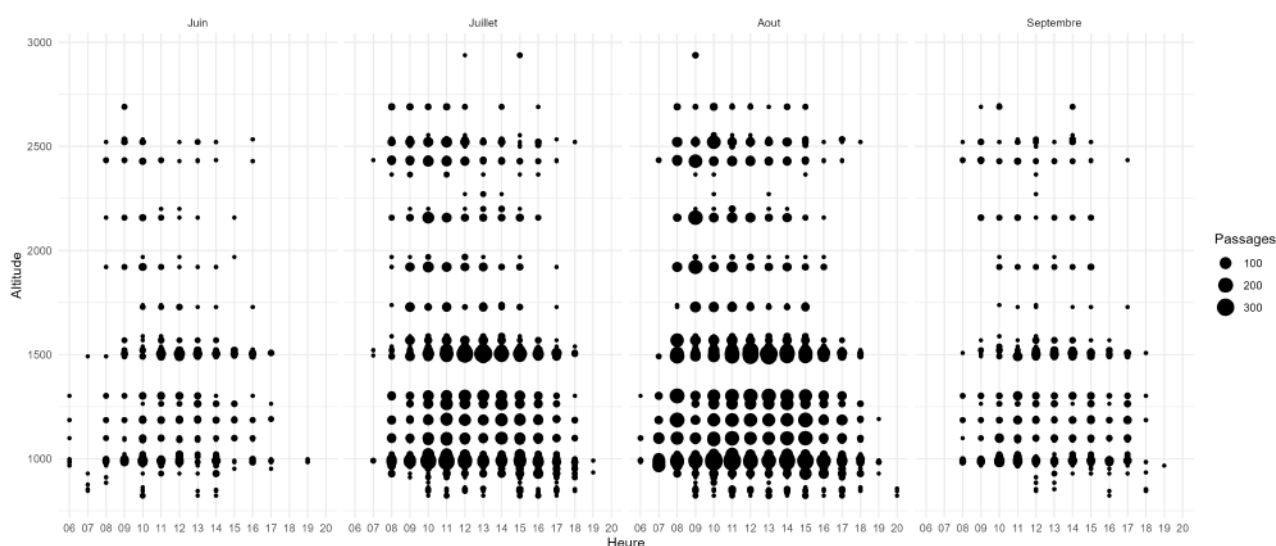


Figure 20 : Distribution temporelle des traces Strava entre Juin et Septembre dans le vallon du Lauvitel en fonction de l'altitude. La taille des points indique le nombre de passages par tronçon de chemin Strava et par heure.

Les pièges photos permettent également de s'intéresser à la diffusion temporelle du passage. Ils permettent notamment de comparer les horaires de passage entre les différents pièges photo. Pour cela, il est important de ne considérer que des périodes où les quatre pièges photo fonctionnent, ce qui restreint drastiquement la période aux données comprises entre le 14/08/2023 et le 25/09/2023. L'analyse de ces données permet par exemple de voir que **le chemin rive gauche est en général privilégié à la montée** (plus de données le matin), tandis que le **chemin rive droite est en général privilégié à la descente** (plus de données l'après-midi ; Figure 21).

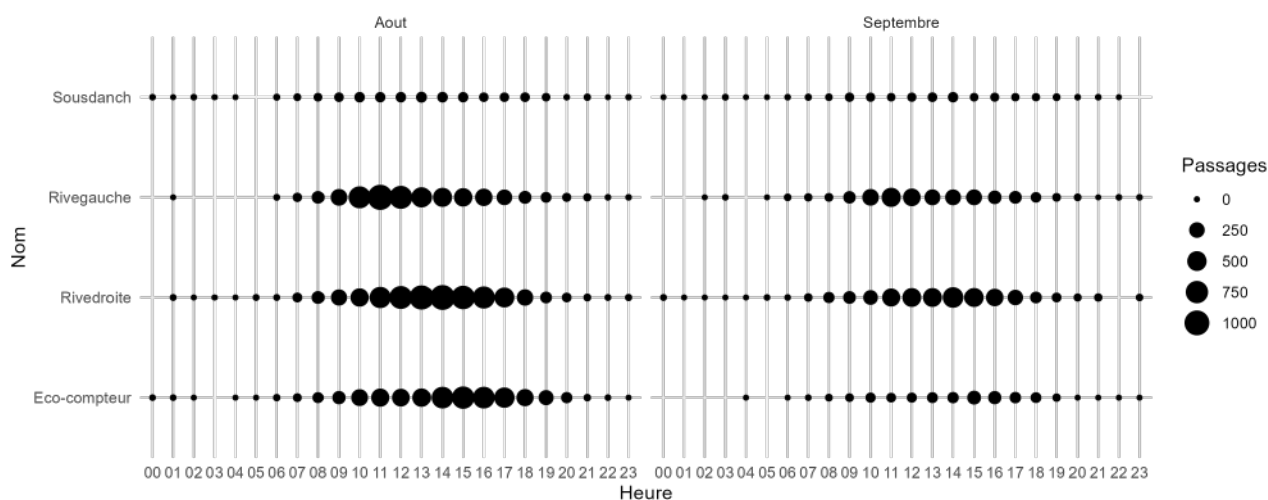


Figure 21 : Distribution horaire des passages devant les pièges photo dans le vallon du Lauvitel.

Plusieurs types de données, à la fois internes et externes, permettent d'étudier en finesse la distribution horaire de la fréquentation du vallon du Lauvitel. Cela permet par exemple de mieux comprendre les activités qui y sont réalisées et d'identifier les horaires où la fréquentation humaine peut avoir le plus d'impact sur la faune présente dans le vallon. On voit ici que la fréquentation s'étale sur l'ensemble de la journée, mais que la finesse des données permet d'identifier quelques variations locales en fonction de l'altitude (avec Strava) ou des différents chemins (avec les pièges photo).

Répartition des différentes activités

Il peut s'avérer utile à l'échelle d'un vallon de mieux comprendre la répartition entre les différentes activités de pleine nature (randonnée, course, cyclisme, ski, escalade...). Quelques données disponibles permettent de caractériser un peu plus finement le type d'activités réalisées sur le territoire du PNE.

Pour étudier cela, on peut par exemple utiliser les données des pièges photos qui nous permettent de distinguer quelques activités ou profils de visiteurs. Les pièges photos permettent notamment d'identifier les passages de canidés (grâce à une identification automatisée, qui ne différencie pas les chiens et les loups). Il est ainsi possible de quantifier le nombre de passages de

chiens, en supposant que l'identification est fiable et constante, en fonction du jour ou de l'heure. D'après les données des quatre pièges photo entre juin et septembre 2023 et 2024, **les canidés représentent 0,47 % des effectifs humains** (donc, en supposant que les identifications ne concernent pas des loups, environ une personne sur 200 serait accompagnée d'un chien). La répartition horaire des canidés sur ces pièges photo est sensiblement identique à celle des passages humains (Figure 22).

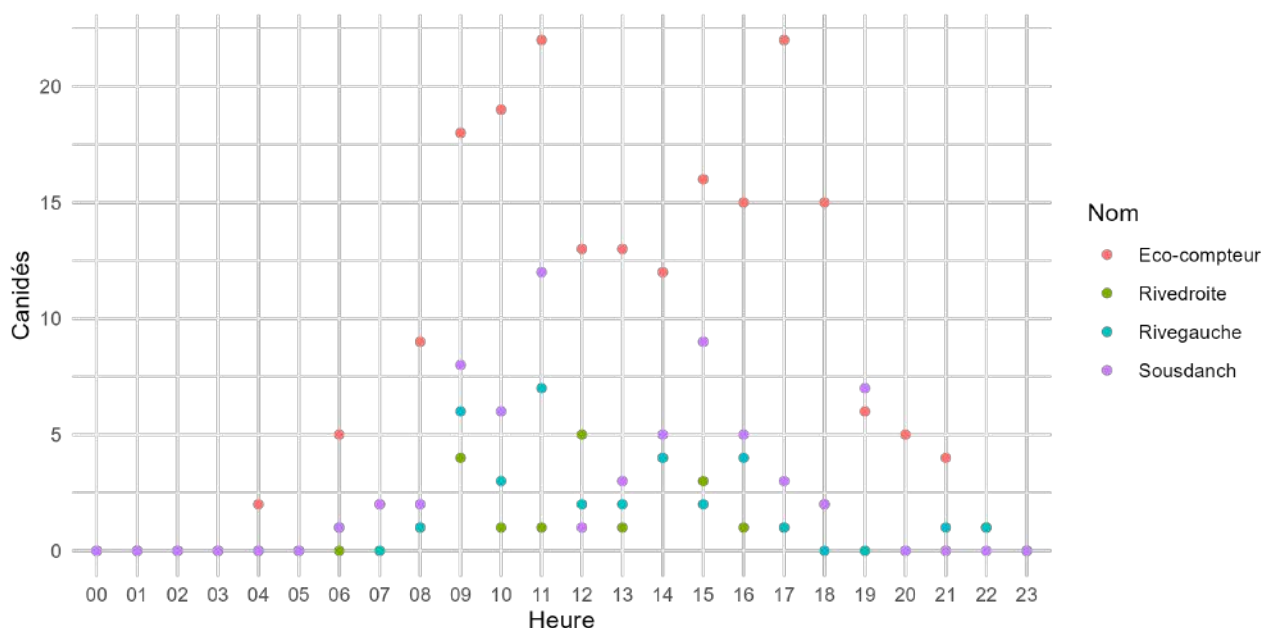


Figure 22 : Répartition horaire des passages de canidés devant les 4 pièges photo du Lauvitel (somme des canidés comptés par heure sur l'ensemble des 188 jours de fonctionnement des pièges photo du Lauvitel entre Juin et Septembre sur les années 2023 et 2024).

Les enquêtes réalisées à la Danchère (726 questionnaires remplis entre 2022 et 2024) permettent également d'en apprendre plus sur les personnes visitant le site. On y découvre que les visiteurs sont principalement des personnes dont la **durée de séjour est de plusieurs jours** (88 % dont 30 % restant au moins 7 jours), qui **logent sur place temporairement** (59 % ayant répondu « village », « station » ou « gîte ») tandis que 16 % sont en itinérance, 15 % sont dans leur résidence principale ou secondaire et 6 % en camping car. La question posée sur le type de randonnée pratiquée permet d'observer que **51 % des visiteurs prévoient une randonnée à la demi-journée ou moins, contre 29 % à la journée, et 20 % qui prévoit de passer la nuit** (les 3/4 en bivouac et 1/4 en refuge).

Autres questions

Même à l'échelle plus fine d'un vallon, il apparaît que les questions auxquelles les données actuelles nous permettent de répondre touchent principalement la répartition spatiale et temporelle de la fréquentation. D'autres questions pourraient cependant être posées à cette échelle si quelques données supplémentaires étaient collectées.

Un des angles morts des données actuelles est l'identification des différentes activités pratiquées dans le parc. Les données dont on dispose regroupent souvent les différentes activités de manière indiscriminée (randonnée, course, vélo) alors qu'il pourrait être intéressant de les étudier séparément. D'autres activités de pleine nature sont elles complètement occultées par les données actuelles, comme le ski de randonnée, l'escalade ou le vol libre. Si ces types de fréquentation sont intéressantes pour le PNE, acquérir des données spécifiques à ces activités s'avérerait utile afin de cartographier les différentes activités présentes sur le parc. Les différentes analyses présentées dans ce rapport montrent la puissance des données basées sur des traces GPS par rapport aux données basées sur des points de comptage ponctuels ; la collecte de traces GPS est alors à favoriser dans la mesure du possible. Les traces GPS de Strava et Outdoorvision pourraient fournir certaines de ces données et permettre une analyse plus fine des types d'activités réalisées dans le vallon. Les pièges photo et les enquêtes permettent également d'affiner notre connaissance sur quelques comportements ciblés, comme le fait de se balader avec un chien ou de passer à vélo. Le travail actuel du PNE s'intéressant aux activités de bivouac paraît particulièrement intéressant et pourrait là-aussi être complété par une récolte de données quantitatives.

Recommandation n°10 : Développer les méthodes de suivi pour des activités de pleine nature autres que la course, la marche et le vélo. Pour l'instant aucune donnée n'est prête à l'emploi sur la fréquentation en ski de randonnée, en escalade, en vol libre, ou encore le bivouac.

De nouveaux dispositifs de suivi pourraient également être mis en place pour s'intéresser à des comportements ciblés qui intéressent le PNE. Si l'on estime que les lacs d'altitude peuvent être menacés par un nombre trop important de baigneurs, on pourrait par exemple mettre en place un dispositif de comptage des comportements de baignade. Il faudrait déployer des pièges photo avec une vue panoramique et qui seraient déclenchés à intervalle régulier (par exemple toutes les minutes). Il pourrait également être intéressant de mesurer si les randonneurs et randonneuses restent sur les sentiers officiels ou s'en écartent beaucoup. Cette question pourrait être étudiée en utilisant les données brutes de Strava ou d'OutdoorVision ou bien, pour une échelle plus fine, avec un déploiement spécifique de pièges photos. Enfin, le déploiement de pièges photo peut permettre d'étudier l'attrait que représentent certains dispositifs de sensibilisation mis en place par le parc (par exemple des panneaux explicatifs) en mesurant la proportion de visiteurs qui s'arrêtent pour consulter ces dispositifs.

Recommandation n°11 : Cibler les comportements que le PNE souhaite étudier (par exemple fréquentation des lacs, sorties de sentiers) et développer des méthodes de suivi adéquates, par exemple par piège photo.

Enfin, il est utile de noter que les données collectées peuvent être mises en relation avec des données externes grâce à leur bon marquage temporel. En effet, on pourrait étudier le lien entre fréquentation et météo en téléchargeant des données météorologiques passées et en les corrélant aux différentes données de fréquentation. De la même façon on pourrait étudier l'impact de certains événements régionaux (un festival, les vacances scolaires, ou les Jeux Olympiques de 2030) sur la fréquentation.

Objectif 3 : utiliser les données de fréquentation pour la protection de la biodiversité

L'espace du PNE combine à la fois une fréquentation touristique élevée, et des espèces sensibles et menacées qui peuvent être perturbées par cette fréquentation. Un des objectifs à long-terme du suivi de la fréquentation est donc l'identification des zones de forte interaction entre la fréquentation touristique et la biodiversité afin de pouvoir gérer les éventuelles tensions. Dans le but d'explorer comment ces données peuvent être utilisées à des fins de protection de la biodiversité, j'ai réalisé une analyse pilote visant à quantifier la fréquentation des zones naturelles sensibles. J'ai appliqué cette analyse pilote à cinq jeux de données variés de biodiversité sensible dans le PNE : les sites de nidification d'aigle royal, les stations botaniques de chardon bleu, les habitats prioritaires Natura2000, les zones humides et les zones de quiétude pour tétras.

Données de biodiversité sensible

Sites de nidification d'aigle royal

Une cinquantaine de couples d'aigles royaux, *Aquila chrysaetos*, espèce charismatique et strictement protégée, se reproduisent sur le territoire du PNE. Or un trop fort dérangement à proximité immédiate de l'aire de nidification peut avoir des conséquences négatives sur l'espèce, telles que l'abandon du nid.

Le PNE suit l'ensemble des aires d'aigle royal, avec un système de rotation d'année en année. Les données des cinq dernières années m'ont été transmises par l'équipe du PNE. J'ai extrait les coordonnées des 336 aires suivies sur ces cinq années, et conservé seulement les aires qui avaient eu une reproduction avérée entre 2021 et 2025 (c'est à dire avec des commentaires contenant les mots « couvaison », « jeune », ou « nourrissage »), pour ne garder que les 64 aires actives.

Stations botaniques de chardon bleu

Le chardon bleu, *Eryngium alpinum*, est également une espèce strictement protégée et à fort enjeu sur le parc puisqu'il est endémique des Alpes et que la plus grosse population mondiale se trouve dans le PNE.

L'équipe du PNE m'a fourni un fichier d'aires de présence de nombreuses plantes, duquel j'ai extrait les points et polygones concernant le chardon bleu et écarté les données trop anciennes (avec une date maximale inférieure à 2010). Les 163 points ont été transformés en cercles d'un mètre de rayon pour permettre un format commun avec les 122 polygones. J'ai ensuite regroupé les données par stations en utilisant des zones tampons de 500m et en ne gardant que les stations incluant au moins 5 données. Au total, ce sont 12 stations de chardon bleu qui ont été identifiées.

Habitats prioritaires Natura2000

Des habitats d'intérêt communautaire ont été définis au niveau européen par la Directive Habitats et parmi eux, des habitats ont été désignés comme prioritaires car plus directement menacés. La cartographie de ces 1538 sites d'habitats prioritaires, couvrant 7 % de la surface du PNE, m'a été envoyée directement par l'équipe du parc et n'a nécessité aucun filtrage supplémentaire. Il est à noter que ce jeu de données n'est pas mis à jour régulièrement, donc peut inclure des zones qui ne sont plus pertinentes.

Zones humides

Un inventaire des zones humides du PNE est en cours dans le cadre du projet ACLIMO (ALCOTRA). La cartographie de ces habitats, couvrant 0,4 % de la surface du PNE, m'a été envoyée directement par l'équipe du parc. Je n'ai gardé ici que les données polygonales de ce fichier, écartant les données de type points (2) ou lignes (2) qui sont incluses dans le geopackage.

Zones de quiétude pour tétras

Des zones protégées pour assurer la quiétude des habitats essentiels au tétras lyre, *Lyrurus tetrix*, ont été mises en place sur le PNE. Elles visent à assurer la quiétude des habitats favorables au tétras durant la saison hivernale. On en compte actuellement 15, se répartissant sur l'ensemble des secteurs du parc. Leur cartographie m'a directement été envoyée par l'équipe du PNE et n'a nécessité aucun filtrage supplémentaire.

Méthodes

Pour chacun des cinq jeux de données de biodiversité sensible présentés ci-dessus, j'ai souhaité comparer la fréquentation humaine entre les lieux qui sont sensibles d'un point de vue de la biodiversité, et ceux qui le sont moins. L'objectif est ainsi de mettre en évidence si ces zones sensibles sont relativement préservées de la fréquentation ou, à l'inverse, les zones de forte interface entre biodiversité et fréquentation humaine.

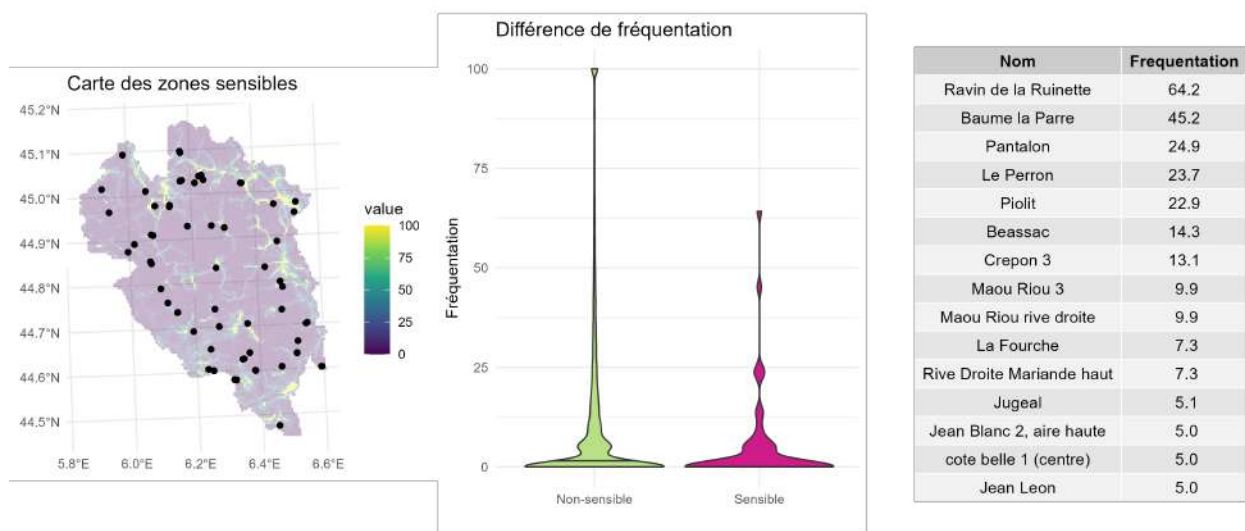
Pour cela, j'ai extrait pour chacune des cellules de la grille du raster de fréquentation si elle inclut des aires d'aigles royaux, des stations de chardon bleu, des habitats prioritaires, des zones humides et des zones de quiétude pour tétras. J'ai ensuite comparé, pour chaque jeu de données, la fréquentation des cellules présentant des zones sensibles à celles n'en présentant pas. Cette comparaison, représentée graphiquement par un `violin_plot` à deux couleurs, est complétée par un test de comparaison de moyennes non-paramétrique de Wilcoxon-Mann-Whitney (car les données de fréquentation ne suivent pas une distribution normale). Un test montrant que la fréquentation est significativement inférieure dans les zones sensibles indique que la biodiversité sensible est relativement préservée de l'impact de la fréquentation humaine. A l'inverse, un test montrant que la fréquentation est significativement supérieure dans les zones sensibles indique que la biodiversité est surexposée à la fréquentation humaine, ce qui est une situation problématique. Entre les deux, un test non significatif indique que la biodiversité n'est pas mise à l'écart de la fréquentation et qu'il est alors possible qu'elle soit impactée.

Dans le but de donner des indications pouvant guider des actions de gestion, la fréquentation moyenne de chaque entité de biodiversité sensible (par exemple chaque aire d'aigle royal ou chaque parcelle d'habitat prioritaire) est également calculée et les entités présentant la plus haute fréquentation sont rapportées sous forme de table. Cela peut aider l'équipe du PNE à identifier les zones où l'interface entre biodiversité sensible et fréquentation humaine est la plus importante et problématique.

Résultats

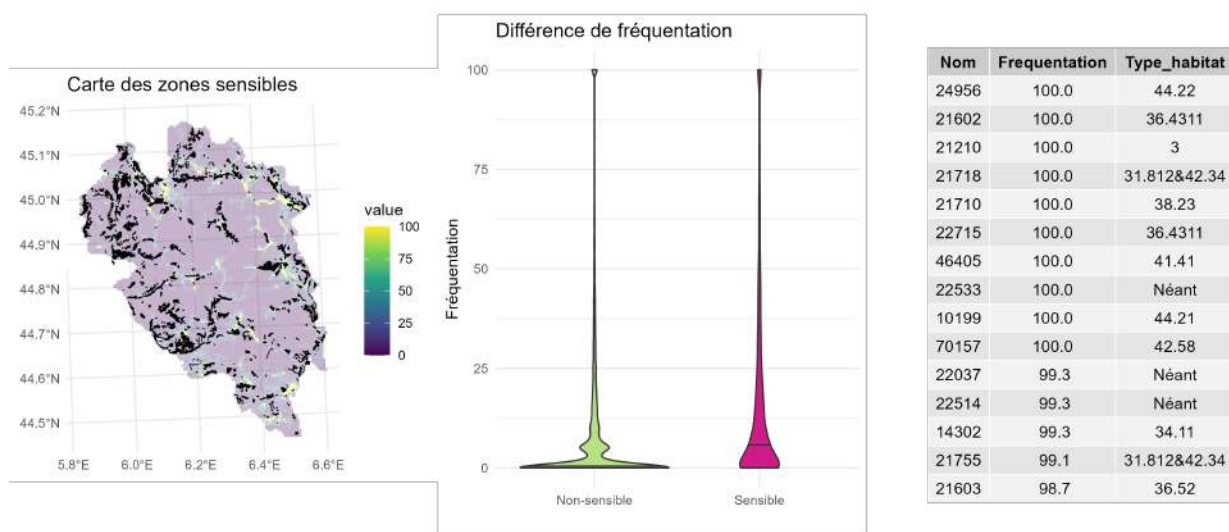
La même analyse a été appliquée aux cinq jeux de données de biodiversité sensible utilisés ici et les résultats sont présentés dans les Figures 22 – 27. Dans les cinq cas, **certaines zones sensibles se trouvent dans des zones de forte fréquentation** comme le montrent les cartes des données brutes superposées à la fréquentation. Pour **trois jeux de données (habitats prioritaires, stations de chardons bleus, zones de quiétude de tétras)**, la fréquentation est **significativement plus élevée dans les zones sensibles que dans les zones non-sensibles**. A l'inverse, elle est significativement plus faible proche des aires d'aigles royaux, et n'est pas significativement différente entre les cellules contenant des zones humides ou non.

De plus, **chaque jeu de données présente des zones sensibles dans des zones à forte fréquentation**. Par exemple, 79 zones humides (Figure 25) et 10 sites d'habitats prioritaires (Figure 24) se trouvent dans les cellules les plus fréquentées (valeur de fréquentation de 100 %). Une attention particulière doit être portée sur ces sites qui présentent une forte fréquentation estivale sur des zones à fort enjeu pour la biodiversité.



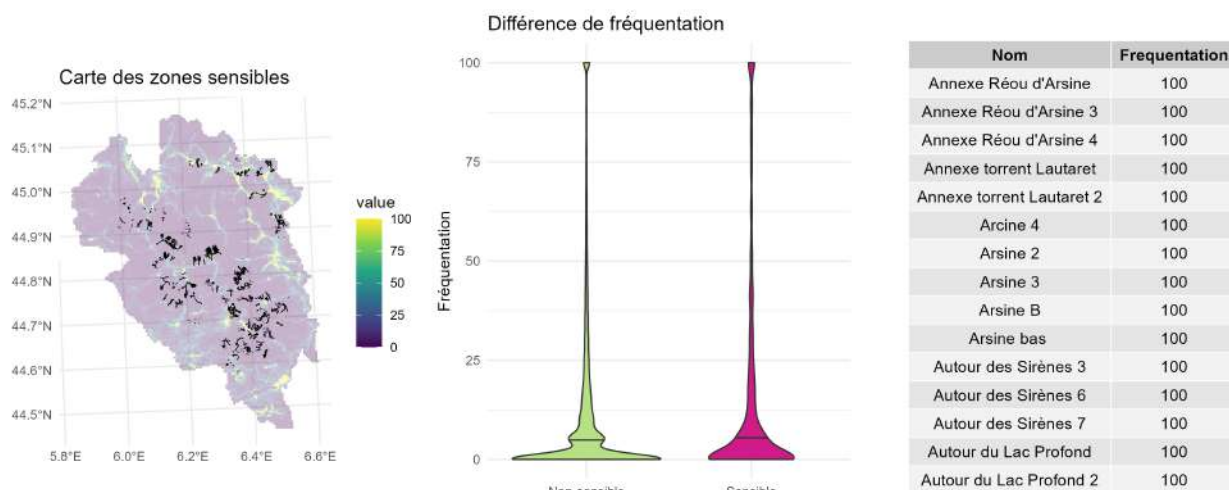
La fréquentation dans les cellules sensibles (5 +/- 11) est en moyenne plus basse que dans les cellules non sensibles (11 +/- 21); P-value=0.0033

Figure 23 : Résultats de l'analyse de la fréquentation standardisée estivale sur les sites d'aires d'Aigle royal ayant une reproduction avérée entre 2021 et 2025 (voir détails dans les Méthodes). Le violon plot (au centre) compare la distribution de la fréquentation entre les zones sensibles et non-sensibles comme un boxplot (le ligne centrale montre la médiane) mais avec plus de détails puisque le violon est plus large aux valeurs les plus fréquentes. Le tableau complet est disponible dans le fichier 'Outputs/Plots/3_Aires d'aigle royal.csv'.



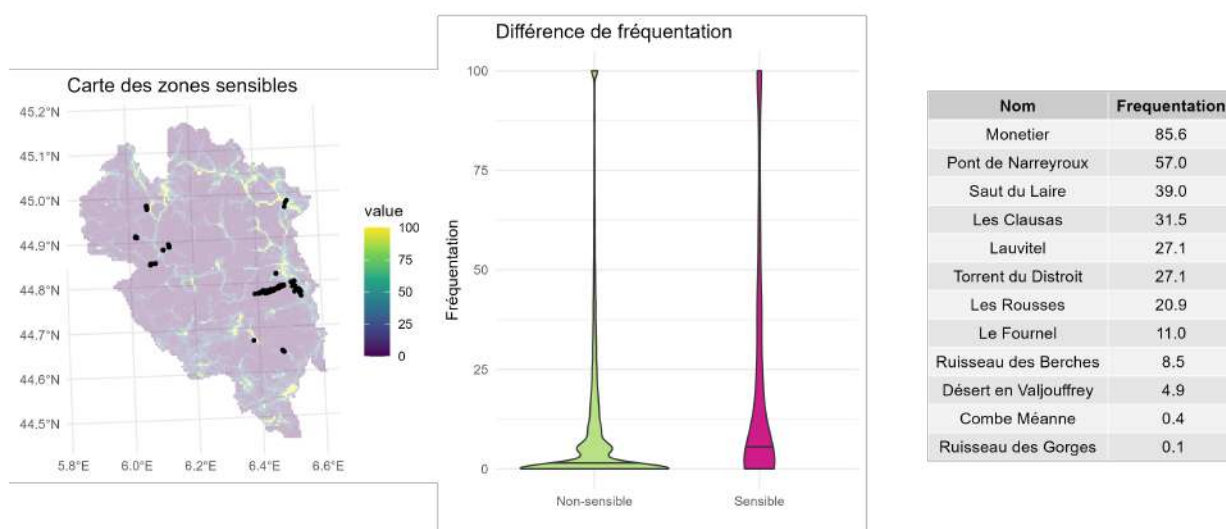
La fréquentation dans les cellules sensibles (16 +/- 24) est en moyenne plus élevée que dans les cellules non sensibles (9 +/- 19); P-value < 10-4

Figure 24 : Résultats de l'analyse de la fréquentation standardisée estivale sur les habitats prioritaires d'intérêt communautaire (voir détails dans les Méthodes). Le nom correspond au code SSP du site et la colonne Type_habitat correspond à la colonne CODE_CB du tableau fourni. Le violon plot (au centre) compare la distribution de la fréquentation entre les zones sensibles et non-sensibles comme un boxplot (le ligne centrale montre la médiane) mais avec plus de détails puisque le violon est plus large aux valeurs les plus fréquentes. Le tableau complet est disponible dans le fichier 'Outputs/Plots/3_Habitats prioritaires.csv'.



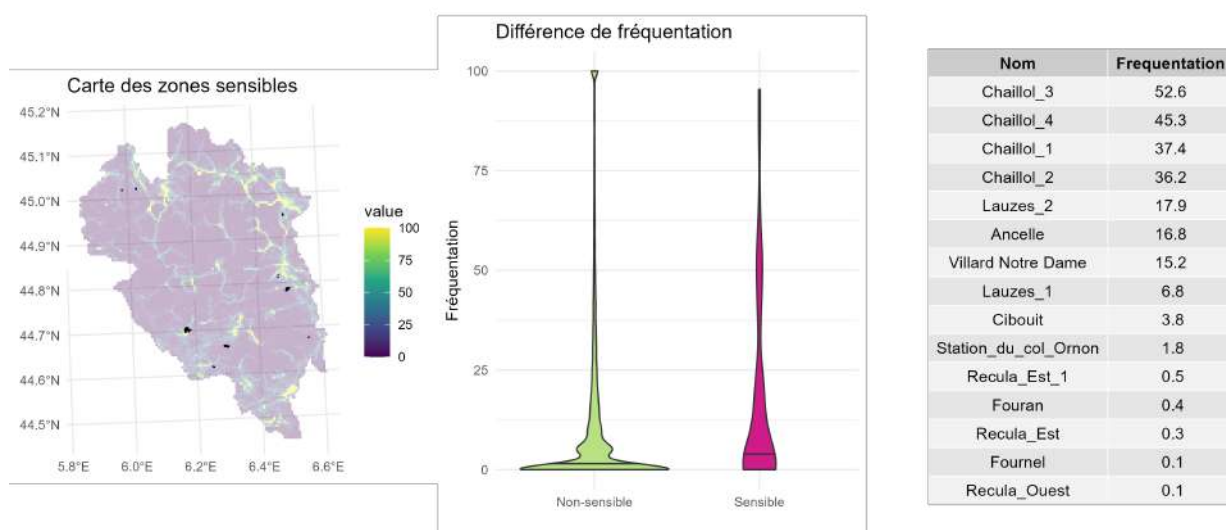
La fréquentation dans les cellules sensibles (14 ± 25) ne diffère pas significativement des cellules non sensibles (11 ± 20); P-value=0.0661

Figure 25 : Résultats de l'analyse de la fréquentation standardisée estivale sur les zones humides (voir détails dans les Méthodes). Le violon plot (au centre) compare la distribution de la fréquentation entre les zones sensibles et non-sensibles comme un boxplot (le ligne centrale montre la médiane) mais avec plus de détails puisque le violon est plus large aux valeurs les plus fréquentes. Le tableau complet est disponible dans le fichier 'Outputs/Plots/3_Zones humides.csv'.



La fréquentation dans les cellules sensibles (20 ± 30) est en moyenne plus élevée que dans les cellules non sensibles (11 ± 20); P-value < 10^{-4}

Figure 26 : Résultats de l'analyse de la fréquentation standardisée estivale sur les stations botaniques de chardon bleu (voir détails dans les Méthodes). Le violon plot (au centre) compare la distribution de la fréquentation entre les zones sensibles et non-sensibles comme un boxplot (le ligne centrale montre la médiane) mais avec plus de détails puisque le violon est plus large aux valeurs les plus fréquentes.



La fréquentation dans les cellules sensibles (16 +/- 24) est en moyenne plus élevée que dans les cellules non sensibles (11 +/- 21); P-value=0.0133

Figure 27 : Résultats de l'analyse de la fréquentation standardisée estivale sur les zones de quiétude pour tétras (voir détails dans les Méthodes). Le violon plot (au centre) compare la distribution de la fréquentation entre les zones sensibles et non-sensibles comme un boxplot (le ligne centrale montre la médiane) mais avec plus de détails puisque le violon est plus large aux valeurs les plus fréquentes. Le tableau complet est disponible dans le fichier 'Outputs/Plots/3_Zones de quiétude pour tetras.csv'.

Discussion

Cette analyse permet de mettre en évidence des zones sensibles d'un point de vue de la biodiversité qui sont fortement fréquentées. En effet, trois des cinq jeux de données de biodiversité sensible testés montrent que les sites sensibles sont plus fréquentés que la moyenne. Pour chacun de ces jeux de données, les sites les plus fréquentés ont été identifiés individuellement (et rapportés dans les tableaux) et appellent à une attention particulière pour voir si les espèces ou habitats concernés sont affectés par cette forte fréquentation, et le cas échéant, si cet impact peut être réduit.

Le choix de l'échelle d'analyse est déterminant dans l'interprétation que l'on peut tirer de telles analyses de corrélation spatiale entre la présence de biodiversité sensible et de fréquentation. En effet, l'analyse présentée ici utilise le raster de 500x500 m qui est pertinent pour comprendre les grandes tendances à l'échelle du parc, mais ne permet pas d'affirmer qu'il y a une vraie perturbation de la biodiversité sensible par la fréquentation. Une fréquentation élevée peut ne pas avoir d'impact sur la biodiversité sensible si la fréquentation ne produit pas de dérangement réel (par exemple un passage important dans une zone de chardon bleu, mais où les randonneurs restent sur les chemins). A l'inverse, une faible fréquentation peut suffire à déranger fortement la biodiversité sensible (par exemple, chez les rapaces, un dérangement unique et ponctuel peut parfois suffire à mener à l'abandon du nid). Des analyses plus spécifiques doivent

donc être menées pour identifier les types de fréquentation particulièrement problématiques, avec une méthode spécifique pour chaque jeu de données. On peut par exemple imaginer utiliser les traces GPS brutes de Strava ou Outdoorvision, ou un dispositif de pièges photo spécifiquement déployés, pour étudier le respect de zones interdites à la fréquentation (comme les zones de quiétude pour tétras en hiver, ou la réserve du Lauvitel). Les traces GPS pourraient aussi être utilisées pour étudier la proximité aux aires d'Aigle royal à laquelle les randonneurs passent. Le choix de l'échelle temporelle est également important : les zones de quiétude de tétras sont principalement pertinentes en hiver, or nos données de fréquentation concernent la période estivale.

Il faut également noter que certaines des zones sensibles utilisées dans cette analyse pilote ont été désignées précisément parce que la biodiversité y était menacée par des activités anthropiques. C'est par exemple le cas des zones de quiétude de tétras, qui définissent des zones importantes pour le tétras dans lesquelles la quiétude des individus n'est pas assurée. Il est alors relativement logique d'observer que la fréquentation y est élevée. C'est également le cas des habitats prioritaires qui sont désignés en fonction des menaces qui pèsent sur eux. Là encore, une analyse à une échelle spatiale plus fine, basée sur les traces GPS brutes, serait alors plus pertinente pour vérifier que la fréquentation à l'intérieur de ces zones de quiétude est plus faible que dans les environs immédiats et ainsi mesurer l'efficacité des zones de quiétude. Il pourrait également être intéressant d'étudier cette question de façon temporelle, pour vérifier que la fréquentation augmente moins dans les zones sensibles.

Recommandation n°12 : Développer des analyses d'interface entre fréquentation et biodiversité sensible à une échelle fine grâce aux traces GPS issues de Strava ou Outdoorvision.

Recommandation n°13 : Déployer des outils de suivi, notamment par piège photo, dans les sites où la biodiversité est particulièrement sensible à la fréquentation afin d'observer les comportements à fort impact sur la biodiversité (par exemple sortie des chemins ou entrée dans des zones interdites).

Discussion générale

Les différentes analyses présentées dans ce rapport montrent que les données de fréquentation dont dispose le PNE sont nombreuses et riches. En effet, la combinaison des données internes et des données externes a permis d'étudier de nombreuses questions sur la fréquentation du territoire du parc, et de produire de nombreuses cartes. Les résultats des différentes analyses sont discutés dans chaque section, mais quelques points généraux de discussion me semblent importants à relever ici.

Sur les données internes, la recommandation la plus importante de ce rapport est d'améliorer le format et la gestion des données. En l'état actuel, il faut plusieurs jours de travail sur les données internes (éco-compteurs, pièges photo, compteurs routiers, maisons de parc, refuges, enquêtes) pour les transformer en un format facilement utilisable pour des analyses. Cela constitue un gros frein à leur utilisation et augmente les chances d'introduction d'erreurs dans ces données.

Lorsque le parc a commencé à collecter les données internes de fréquentation, il n'existait pas de source externe de données tels qu'il en existe maintenant. En effet, plusieurs jeux de données sont disponibles aujourd'hui, parfois gratuitement, pour cartographier la fréquentation de manière précise (Strava, Outdoorvision, iNaturalist, PlantNet). Il est donc important de réfléchir au rôle des données internes, pour qu'il soit en complément et pas en doublon des jeux de données externes. Les données externes étant basées sur des traces ou points GPS sur l'ensemble du territoire, elles offrent une couverture du parc et une finesse spatiale que les données internes ne pourront jamais égaler. En revanche, les données internes présentent trois grands avantages par rapport aux données externes. D'abord, elles permettent de mesurer des tendances temporelles d'une année sur l'autre (ce qui est difficile avec les données externes puisque l'utilisation des applications telles que Strava ou iNaturalist est récente et très dynamique). Ensuite, les données internes permettent d'accéder à des abondances absolues de fréquentation (par exemple un nombre exact de passages avec les éco-compteurs, là où les données externes ne nous informent que sur les personnes utilisant telle ou telle application), cela peut s'avérer utile pour quantifier de manière non relative la fréquentation, ou pour calibrer les données externes (comme cela a été fait pour la cartographie standardisée de fréquentation du parc). Enfin, les données internes peuvent permettre d'étudier des comportements ciblés d'intérêt pour le PNE, notamment par le biais des pièges photo, tels que le passage de chiens, la baignade, le bivouac ou le respect des sentiers. A l'avenir, il paraît important d'orienter la collecte de ces données internes pour maximiser ces avantages de manière plus ciblée, afin qu'elles soient plus complémentaires avec les données externes.

Avec ces différentes données nous avons pu répondre à trois grands besoins du parc. D'abord, la production d'information sur la fréquentation générale du parc avec la carte interactive, la carte standardisée de fréquentation, l'identification des zones blanches de suivi interne, et les analyses temporelles. Ensuite, à une échelle plus fine (celle du vallon du Lauvitel), nous avons pu étudier la dispersion des randonneurs le long du vallon et hors des sentiers, la répartition de la fréquentation au cours des heures de la journée, ainsi que la répartition de certaines activités. Enfin, nous avons

pu croiser la carte de fréquentation avec la distribution des zones sensibles de biodiversité afin d'identifier les zones de forte interaction qui pourraient poser problème. Les méthodes développées pour chacune de ces questions peuvent sans doute être améliorées, mais proposent un point de départ pour des études plus poussées sur des points particuliers.

D'autres questions pourraient être étudiées à l'avenir si certaines données étaient récoltées ou partagées. Le plus grand manque actuel dans les données est la fréquentation hivernale pour laquelle nous avons très peu de données. La façon la plus évidente de combler ce manque serait d'accéder aux données hivernales de Strava et Outdoorvision (notamment ski de randonnée), ce qui permettrait de créer une carte de fréquentation standardisée focalisée sur l'hiver. D'autres questions pourraient également être étudiées s'il était possible de séparer les différentes activités dans les données internes et externes (course, randonnée, vélo...) afin de pouvoir cartographier plus finement ces différentes activités. Enfin, les pièges photo peuvent permettre d'étudier des comportements ciblés s'ils sont déployés spécifiquement pour les étudier : par exemple pour quantifier et cartographier les baignades dans les lacs du parc, ou la pénétration de zones sensibles.

En conclusion, je pense que le PNE a à sa disposition de nombreuses données qui permettent d'établir un bon état des lieux de la fréquentation du parc. Ces données sont déjà une très bonne base et le PNE peut être fier de la qualité et la couverture de ces données. Elles pourront bien sûr être complétées par d'autres données, lorsque des questions spécifiques se poseront (par exemple pour mieux connaître la fréquentation en haute altitude, pour cartographier les activités d'hiver, ou pour s'intéresser à un comportement particulier). En attendant, la priorité doit être, à mon avis, de faciliter l'utilisation de ces données, et la mise à jour régulière des données internes et externes ainsi que des produits de ces analyses (notamment la carte interactive) pour valoriser au mieux les données existantes.

Livrables supplémentaires

En complément de ce rapport, **plusieurs livrables** ont été fournis au PNE :

- une synthèse des principaux résultats (intégrée au début de ce rapport mais également fournie dans un document indépendant)
- deux cartes interactives au format HTML (une complète et une excluant les données brutes de Strava et Outdoorvision, pour une navigation plus rapide)
- le raster de fréquentation estivale standardisée au format TIFF
- l'ensemble des scripts développés dans cette prestation, notamment le script '0.BTA fonctions preparation donnees.R' incluant les fonctions développées pour le PNE et détaillées ci-dessous

Fonctions R développées pour le PNE (voir script '0.BTA fonctions preparation donnees.R') :

BTA_PreparerRoutier : Fonction pour préparer les données de compteurs routiers. Vu l'hétérogénéité des données brutes de compteurs routiers, cette fonction est assez complexe et spécifique aux fichiers déjà existants pour les compteurs routiers. Si le format des tableaux et des noms de fichiers évoluent dans les prochaines années, cette fonction risque de ne pas fonctionner voire de causer des erreurs; à utiliser avec prudence.

BTA_Telecharge_iNaturalist et **BTA_Telecharge_PlantNet** : Fonction pour télécharger et formater les données iNaturalist / PlantNet. Très facile d'utilisation, il suffit de donner le shapefile des secteurs du parc en argument, pour mettre à jour les données à la fréquence souhaitée.

BTA_PrepareStrava : Fonction pour préparer les données flux de Strava. Cette fonction lit les fichiers qui m'ont été envoyés par le PNE et les découpe selon la grille du raster de 500x500m. Elle devra être adaptée si les fichiers d'entrée sont changés (notamment si de nouvelles années sont ajoutées).

D'autres fonctions sont incluses dans ce script, notamment pour créer les fenêtres popup de la carte interactive. D'autres fonctions sont également incluses dans les différents scripts d'analyse (par exemple pour l'analyse de biodiversité sensible), afin de systématiser les analyses qui y sont présentées (voir les détails dans les scripts concernés).

Références

- Affouard A., Joly A., Lombardo J., Champ J., Goeau H., Chouet M., Gresse H., Bonnet P. (2025). PI@ntNet observations. Version 1.9. PI@ntNet. Occurrence dataset <https://doi.org/10.15468/gtebaa> accessed via rgbif on 2025-06-19.
- Barve V., Hart E. (2025). rinat: Access 'iNaturalist' Data Through APIs. R package version 0.1.10, <https://doi.org/10.32614/CRAN.package.rinat>
- Bates D., Maechler M., Bolker B., Walker S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1) : 1-48. doi:10.18637/jss.v067.i01.
- Chamberlain S., Barve V., Mcglinn D., Oldoni D., Desmet P., Geffert L., Ram K. (2025). rgbif: Interface to the Global Biodiversity Information Facility API. R package version 3.8.2, <https://CRAN.R-project.org/package=rgbif>.
- Fong Y., Huang Y., Gilbert P., Permar S. (2017). chngpt: threshold regression model estimation and inference. *BMC Bioinformatics* DOI:10.1186/s12859-017-1863-x.
- IGN, OFB (2025). Bassins versant topographiques - Métropole 2023 - BD Topage®. Téléchargée de <https://www.data.gouv.fr/datasets/bassins-versant-topographiques-metropole-2023-bd-topage-r/informations> le 5 Juin 2025.
- Matthews TJ., Triantis K., Whittaker RJ., Guilhaumon F. (2019). sars: an R package for fitting, evaluating and comparing species–area relationship models. *Ecography*, **42**: 1446-1455. doi:10.1111/ecog.04271
- Miller J. (2010). Species distribution modeling. *Geography Compass*, 4: 490–509.
- Mostert P.S. and O'Hara R.B. (2023). PointedSDMs: An R package to help facilitate the construction of integrated species distribution models. *Methods in Ecology and Evolution*, 14(5): 1200–1207. <https://doi.org/10.1111/2041-210X.14091>
- R Core Team (2025). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org>

Annexes

Annexe 1 : Nombre de dispositifs de suivi (Eco-compteurs, compteurs routiers, pièges photo, refuges, maisons de parc ; somme dans la colonne « Suivi ») et fréquentation moyenne des différents chemins de randonnée du PNE.

Chemin	Suivi	Eco-compteurs	Compteurs routier	Pièges photo	Refuges	Maisons de parc	Fréquentation
Tour Claree-Ecrins en 6 jours	5	3	0	1	0		1 100
Lac Lauvitel	5	1	0	4	0		0 76.6
Tour des refuges en Vallouise	5	1	0	0	3		1 65.9
GR54 - Muzelle-Lauvitel	5	1	0	4	0		0 40.8
Tour des Refuges en Valgaudemar 4 jours	5	2	0	0	3		0 22.3
Tour des Refuges en Valgaudemar 7 jours	5	2	0	0	3		0 20.9
GR54 - Valgaudemar	5	1	0	0	3		1 14.8
Refuge du glacier Blanc	4	1	0	0	2		1 100
Refuge Alpe de Villar d'Arene	4	2	0	1	1		0 62.5
Refuge Adele Planchard	4	2	0	1	1		0 46.1
Tour du Vieux Chaillol	4	1	0	0	3		0 16
Tour du Combeynot en 3 jours	3	2	0	0	0		1 100
Tour du Combeynot en 4 jours	3	2	0	0	0		1 97.6
Lacs d'Arsine	3	2	0	0	0		1 94.6
Glacier noir	3	1	0	0	1		1 73
GR54 - Romanche	3	2	0	1	0		0 61.9
Breche du Perier	3	1	0	2	0		0 61.2
Tour de l'Aiguille de Venosc	3	1	0	2	0		0 56.8
Refuge du Pave	3	2	0	1	0		0 40.1
Refuge du Tourond	3	1	0	2	0		0 36.6
Refuge Temple Ecrins	3	1	0	0	2		0 36.3
Chabourneou-Vallonpierre	3	1	0	0	2		0 26.1
Refuge de l'Aigle	3	1	0	1	1		0 21.8
Tour du Vieux Chaillol en 6 jours	3	1	0	1	1		0 14.6
Lac de la Douche	2	1	0	0	0		1 100
Refuge du Carrelet	2	1	0	0	1		0 41.5
Col des Terres Blanches	2	2	0	0	0		0 37.9
Col de Vallonpierre	2	1	0	0	1		0 35.5

Boucle du Pigeonnier	2	1	0	0	1	0 32.8
Refuge du Chatelleret	2	0	0	0	2	0 29.1
Vallon de Charges	2	1	0	0	0	1 17
GR54 - Monetier	1	1	0	0	0	0 91.4
Grand Tour des Ecrins VTT - Monetier	1	0	0	0	0	1 82.7
Casset-Lauzet	1	0	0	0	0	1 74.8
Lac Noir et lac Lerie	1	1	0	0	0	0 60.7
Sentier des Crevasses	1	1	0	0	0	0 57.7
Grand Tour des Ecrins VTT - Romanche	1	1	0	0	0	0 52.8
Saut du Laire	1	1	0	0	0	0 50.5
Sentier des mules	1	1	0	0	0	0 46.7
Lac Lauzon	1	1	0	0	0	0 43.9
Refuge du Pelvoux	1	0	0	0	1	0 43.6
Tour de La Grave	1	1	0	0	0	0 42.1
GR54 - Emparis	1	1	0	0	0	0 39.8
Tour du Signal de La Grave	1	1	0	0	0	0 39
Boucle Dormillouse	1	1	0	0	0	0 38.9
Refuge des Bans	1	0	0	0	1	0 38.4
Cascade de la Pissarotte	1	0	0	0	0	1 30.6
Tour du Grand Pinier	1	1	0	0	0	0 30.2
Lacs Dormillouse	1	1	0	0	0	0 29.9
Grand tour des Ecrins VTT - Col d'Ornon	1	0	0	0	0	1 28
Tour du Plateau d'Emparis en 3 jours	1	1	0	0	0	0 27.9
Lacs de Petarels	1	1	0	0	0	0 27.2
Lac Lovitel	1	1	0	0	0	0 26.3
Lac Lautier	1	0	0	0	1	0 25.7
Tour des Refuges en Valgaudemar en 7 jours	1	0	0	0	1	0 25.7
GR54 - Champoleon	1	0	0	0	1	0 24.4
Col de Freissinieres	1	1	0	0	0	0 22.2
Col du Cheval de bois	1	0	0	0	1	0 20.5
Boucle de l'Oussella	1	0	0	0	0	1 19.5
Lac du Distroit	1	1	0	0	0	0 18.7
Refuge du Pre de la Chaumette	1	0	0	0	1	0 18.2
Grand Tour des Ecrins VTT - Valsenestre	1	0	0	0	0	1 15.7
Lac Gary	1	0	0	0	0	1 14
Tour du Plateau d'Emparis en 4 jours	1	1	0	0	0	0 13.7
GR54 - Valsenestre	1	1	0	0	0	0 13.2
Refuge Font Turbat	1	0	0	0	1	0 12.7
Tour du Vieux Chaillol en	1	0	0	0	1	0 12.1

trail						
Jas des Agneaux	1	0	0	0	0	1 11.7
Tour du Pic de Valsenestre	1	0	0	0	1	0 11.4
Grand Tour des Ecrins VTT	1	1	0	0	0	0 10.9
Grand Tour des Ecrins VTT - Le Rabioux	1	0	0	0	0	1 8.8
La Lavey	1	0	0	0	1	0 6.8
Tour Gourmand	1	0	0	0	1	0 0.2
Tour du Taillefer en 6 jours	1	1	0	0	0	0 0
GR54 - Col d'Ornon	0	0	0	0	0	0 64.4
Tour du Pays des Ecrins en 6 jours	0	0	0	0	0	0 63.4
Tour du Pays des Ecrins en 7 jours	0	0	0	0	0	0 63.1
Cascade de la Pissette	0	0	0	0	0	0 59.8
Lac de l'Eychauda	0	0	0	0	0	0 57.6
Lac Muzelle	0	0	0	0	0	0 56.1
Grand Tour des Ecrins VTT - Emparis	0	0	0	0	0	0 55.9
Belvedere des Tetes	0	0	0	0	0	0 55
Lacs du Plateau d'Emparis	0	0	0	0	0	0 53.3
Refuge du Sele	0	0	0	0	0	0 49.2
Pic de Morgon	0	0	0	0	0	0 49
GR54 - Vallouise-Pelvoux	0	0	0	0	0	0 48.2
Croix Saint-Philippe	0	0	0	0	0	0 45
Tour des Aiguilles de Chabrières	0	0	0	0	0	0 43.6
Tour du lac de Serre-Poncon a VTT	0	0	0	0	0	0 43.1
Sentier des Pyramides	0	0	0	0	0	0 42.2
Plateau du Taillefer	0	0	0	0	0	0 41.3
Lac du Goleon	0	0	0	0	0	0 41.2
Bosse de Clapouse	0	0	0	0	0	0 40.6
GR54 - Dormillouse	0	0	0	0	0	0 40.2
Oulles du Diable	0	0	0	0	0	0 38.7
Lac des Pisses	0	0	0	0	0	0 38.6
Tour du Pays des Ecrins en 3 jours	0	0	0	0	0	0 38.3
Les Sagnières	0	0	0	0	0	0 36.5
Lacs d'Orcières	0	0	0	0	0	0 36.4
Mont Guillaume	0	0	0	0	0	0 36.4
Sentier des lutins	0	0	0	0	0	0 36.4
La Pinouse	0	0	0	0	0	0 36.2
Vallon de Navette	0	0	0	0	0	0 35.4
Col Laurichard	0	0	0	0	0	0 35.2
La tete d'Oreac	0	0	0	0	0	0 34.8

Le balcon du Champsaur	0	0	0	0	0	0 34.7
GR54 - Clavans	0	0	0	0	0	0 32
Sentier des canaux	0	0	0	0	0	0 30.8
Cabane de Jas Lacroix	0	0	0	0	0	0 30.3
Cabane des Grands Plans	0	0	0	0	0	0 30.3
Fort de Reallon	0	0	0	0	0	0 29.9
Boucle de Tiriere	0	0	0	0	0	0 29.7
Lac de Combeynot	0	0	0	0	0	0 29.7
Les Plagnes	0	0	0	0	0	0 28.7
Cascade de la Pisse	0	0	0	0	0	0 28.4
Lac des Quirlies	0	0	0	0	0	0 28.1
Col de Cote Belle	0	0	0	0	0	0 27.9
Col de la Coupa	0	0	0	0	0	0 27.5
Grand Tour des Ecrins VTT - Reallon	0	0	0	0	0	0 26.9
Promontoire de Pregentil	0	0	0	0	0	0 25.5
Le refuge du Pre de la Chaumette	0	0	0	0	0	0 24.8
Col de l'Aup Martin	0	0	0	0	0	0 24.1
Lac de Puy Aillaud	0	0	0	0	0	0 24
Refuge Soreiller	0	0	0	0	0	0 23.8
Chardon bleus	0	0	0	0	0	0 23.7
GR54 - La Bonne	0	0	0	0	0	0 23.3
Le Palastre	0	0	0	0	0	0 23
Bouchier et le Clot de la Siva	0	0	0	0	0	0 22.6
Lac Labarre	0	0	0	0	0	0 22.5
Combe Oursiere	0	0	0	0	0	0 22.2
Lacs de Crupillouse	0	0	0	0	0	0 21.9
Grand Tour des Ecrins VTT - Severaissette	0	0	0	0	0	0 20.9
Cuchon/Rouanne	0	0	0	0	0	0 20.5
Hameau des Peines	0	0	0	0	0	0 20.1
Vallon de Valsenestre	0	0	0	0	0	0 20.1
Grand Tour des Ecrins VTT - Dormillouse	0	0	0	0	0	0 19.9
Tete de Vautisse	0	0	0	0	0	0 19.2
Le sentier du Roy	0	0	0	0	0	0 19
Grand tour des Ecrins VTT - Vallouise-Pelvoux	0	0	0	0	0	0 18.8
Pas de l'Olan	0	0	0	0	0	0 18.5
Belvedere Roche Rousse	0	0	0	0	0	0 18.4
Grand Tour des Ecrins VTT - Pont-du-Fosse	0	0	0	0	0	0 18
Pic du Mas de La Grave	0	0	0	0	0	0 17.4
GR54 - Malsanne	0	0	0	0	0	0 16.8
6 cabanes	0	0	0	0	0	0 16.7

Lac du Vallon	0	0	0	0	0	0 16.6
Col des Tourettes	0	0	0	0	0	0 16.5
Roche du Touret	0	0	0	0	0	0 16.5
Vallon de la Chalpe	0	0	0	0	0	0 16.2
Grand Tour des Ecrins VTT - Clavans	0	0	0	0	0	0 15.9
Refuge de la Selle	0	0	0	0	0	0 15.9
Cabane de Rif Meyol	0	0	0	0	0	0 15.7
Pas de la Cavale	0	0	0	0	0	0 15.4
Travers des lacs	0	0	0	0	0	0 15
Tour du Taillefer en 4 jours	0	0	0	0	0	0 14.8
Col de Font Froide	0	0	0	0	0	0 14.7
Circuit de Clot Laffont	0	0	0	0	0	0 14.6
Crêtes de la croix du Carrelet	0	0	0	0	0	0 14.2
Vallon du Couleau	0	0	0	0	0	0 13.2
Au-dessus des Drac	0	0	0	0	0	0 11.8
Grand Tour des Ecrins VTT - Valgaudemar	0	0	0	0	0	0 10.6
Tour du Champsaur VTT 5 jours	0	0	0	0	0	0 7
Tour du Champsaur VTT 3 jours	0	0	0	0	0	0 1.4
Grand Tour des Ecrins - Taillefer	0	0	0	0	0	0 NA
Meollion	0	0	0	0	0	0 NA