



**HAL**  
open science

# Estimation de la fréquentation de tourisme et de loisirs des sites naturels d'intérêt patrimonial

Jacques Veslot

► **To cite this version:**

Jacques Veslot. Estimation de la fréquentation de tourisme et de loisirs des sites naturels d'intérêt patrimonial. Inrae. 2024. hal-04668906

**HAL Id: hal-04668906**

**<https://hal.science/hal-04668906v1>**

Submitted on 7 Aug 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0  
International License

## Projet Valo-Sites

Évaluation des effets socio-économiques  
des politiques de gestion et valorisation de sites naturels

Tâches T1-A:

**Estimation de la fréquentation de tourisme et de loisirs  
des sites naturels d'intérêt patrimonial**

Exemple du site Chaîne des Puys-Faille de Limagne

***Rapport final (avril 2022)***

Jacques VESLOT

## TABLE DES MATIERES

1.	Introduction .....	4
1.1.	Enjeux et problématique .....	4
1.2.	Quelle fréquentation ? .....	5
1.3.	Méthodes de comptage .....	6
1.4.	Projet.....	9
1.5.	Zone d'étude .....	10
1.6.	Objectifs.....	16
1.7.	Plan .....	17
2.	Estimation à partir des sites dénombrables .....	18
2.1.	Données disponibles .....	18
2.1.1.	Éco-compteurs.....	19
2.1.2.	Panoramique des Dômes .....	29
2.1.3.	Sites payants.....	30
2.1.4.	Aydat .....	34
2.1.5.	Autres données .....	34
2.2.	Traitement des données d'éco-compteurs .....	38
2.2.1.	Diagnostic des séries .....	38
2.2.2.	Traitements.....	51
2.2.3.	Combinaison et extrapolation .....	53
2.3.	Autres sources.....	55
	Panoramique .....	55
	Billetteries.....	55
	Aydat.....	56
2.4.	Estimation globale .....	58
2.4.1.	Agrégation des sources .....	58
2.4.2.	Conversion des visites en visiteurs .....	59
3.	Méthodes alternatives.....	60
3.1.	Comptages routiers .....	60
3.1.1.	Expérimentation .....	60
3.1.2.	Comptages vidéo.....	71
3.2.	Flux Vision .....	73
3.2.1.	Pourquoi tester le service Flux Vision d'Orange ? .....	73
3.2.2.	Présentation de Flux Vision .....	75
3.2.3.	Zones d'étude .....	80
3.2.4.	Deux offres Flux Vision.....	84

3.2.5. Diagnostic des données .....	87
3.2.6. Corrections .....	90
4. comparaisonS des méthodes .....	95
4.1. Comparaison Flux Vision / agrégat des sites dénombrables .....	95
4.1.1. Sous-zone Lemptégy-Vulcania .....	95
4.2.2. Zone complète réduite .....	98
4.2. Autres données Flux Vision .....	100
4.2.1. Comparaison avec les données d'enquête .....	100
4.2.2. Autres informations Flux Vision .....	102
4.2.3. Synthèse des résultats sur Flux Vision .....	106
5. Conclusion .....	107
5.1. Avantages / Inconvénients des différentes approches .....	108
5.2. Critères de choix .....	109
5.3. Perspectives .....	110
Bibliographie .....	111
Liste des figures .....	115

## LISTE DES ABREVIATIONS

Abréviation	Signification
<b>ADT</b>	Agence de Tourisme
<b>CDPFL</b>	Chaîne des Puys – Faille de Limagne
<b>CRT</b>	Comité Régional de Tourisme
<b>EC</b>	Éco-compteurs
<b>FV</b>	Flux Vision
<b>FVT</b>	Flux Vision Tourisme
<b>GMK</b>	Flux Vision Géomarketing
<b>PNRVA</b>	Parc Naturel Régional des Volcans d'Auvergne
<b>PDD</b>	Panoramique des Dômes
<b>SD</b>	Agrégat des fréquentations sur sites dénombrables

## 1. INTRODUCTION

Depuis les années 50, on observe une augmentation continue du tourisme en France (Dauphin et al., 2008) et, au cours des dernières années, une progression notable dans les zones de montagne, les zones rurales et les zones de moyenne montagne (Gidrol, 2019), notamment liée à la présence de sites naturels et, en particulier, des sites classés, « qu'il faut impérativement avoir vus ».

Dans les sites naturels et protégés, le problème de la gestion des flux touristiques se pose à la fois sur le plan de leurs retombées économiques et sociales (valorisation) et des impacts environnementaux (protection). Le besoin de suivre la fréquentation répond à ce double objectif de gestion de la pression touristique et de valorisation essentielle à l'économie touristique car elle constitue un levier de développement local, en termes d'emplois et de revenus.

A la pression directe de la fréquentation touristique s'ajoute celle de la fréquentation résidentielle, en forte croissance ces dernières années, accrue par l'attractivité du territoire, ainsi que la pression indirecte que constitue l'ensemble des pressions anthropiques liées à l'attractivité de la zone, notamment liées à l'accroissement de la population (artificialisation des milieux, pollutions, introduction d'espèces invasives...).

Le site Chaîne des Puys – Faille de Limagne, qui fait partie du Parc National des Volcans d'Auvergne (PNRVA) depuis sa création en 1977, fait l'objet de politiques de gestion menées par le conseil départemental du Puy-de-Dôme en collaboration avec le PNRVA, les collectivités et les services de l'état, politiques qui se sont notamment traduites par la labellisation « Grand site de France » du Puy de Dôme en 2008, par l'aménagement du sommet et l'installation du train à crémaillère en 2014. L'inscription du site CDPFL, le 2 juillet 2018, sur la liste du patrimoine mondial de l'UNESCO, en tant que haut lieu tectonique, implique encore de nouveaux objectifs de protection (Lefebvre & Moncorps, 2010).

De ce fait, le suivi de la fréquentation des sites naturels constitue un élément essentiel des politiques de gestion, qui sous-tend à la fois les problématiques de protection des milieux et de développement des territoires.

### 1.1. ENJEUX ET PROBLEMATIQUE

Parmi les principaux enjeux potentiels qui pourraient orienter le choix de la méthode de quantification de la fréquentation, on peut citer :

- La nécessité d'une méthode reproductible permettant notamment des comparaisons interannuelles sur des indicateurs d'intérêt ;
- Le besoin d'un comptage exhaustif, et représentatif des composantes de la fréquentation, notamment pour les évaluations économiques ;
- L'importance de ces mesures pour calibrer des enquêtes, ainsi que pour définir des quotas ou réaliser des redressements ;
- Si possible, avoir des premiers éléments de qualification de la fréquentation : par exemple la distinction touristes/excursionnistes ;
- Enfin, idéalement, avoir des mesures plus fines dans le temps et dans l'espace, pour d'une part pouvoir faire des distinctions par site ou secteur, mais surtout avec une

résolution temporelle plus fine (par exemple journalière), pour pouvoir modéliser et ainsi prédire, voire expliquer les évolutions de la fréquentation.

Pour connaître l'importance réelle des flux de visiteurs, il est important d'obtenir des estimations les plus exhaustives possibles, sans omettre des composantes de la fréquentation, notamment pour le calcul des indicateurs liés aux impacts économiques (revenus, emplois...), sociaux (nuisances...) et environnementaux (pression exercée sur les milieux, les habitats...).

Pour le suivi de cette fréquentation, une approche reproductible, dans le sens de *répliquable*, c'est-à-dire capable de produire, d'une année sur l'autre, avec des personnes pouvant être différentes, une estimation portant sur le même objet. En outre, pour assurer cette reproductibilité, il est nécessaire de se limiter à des approches qui ne soient pas trop techniques d'une part, et dont le coût reste raisonnable d'autre part.

Outre la connaissance de la fréquentation globale, ce suivi est essentiel pour calibrer et planifier les enquêtes portant sur les visiteurs (plans d'échantillonnage, redressements), enquêtes elles-mêmes nécessaires à la qualification et à la gestion de la fréquentation, par exemple pour l'évaluation des retombées économiques.

Pour cet usage notamment, il est important de ne pas se limiter à la seule connaissance des cumuls annuels de fréquentation, mais de connaître la répartition de cette fréquentation dans le temps et dans l'espace. Une connaissance fine de la fréquentation, à un grain infra-annuel, permettra une analyse des variations saisonnières, la modélisation des flux, permettant de prédire, d'imputer, voire d'expliquer certains phénomènes.

Enfin, idéalement, certaines méthodes de suivi quantitatif de la fréquentation peuvent apporter également des premiers éléments de qualification de la fréquentation, par exemple sur l'origine et sur le type des visiteurs (touristes, excursionnistes, résidents...), voire sur leurs profils (données socio-démographiques...).

## 1.2. QUELLE FREQUENTATION ?

On peut distinguer les visiteurs selon leur type en :

- Touristes (visiteurs avec nuitée hors domicile), logeant ou non dans le site, que l'on peut classer selon le type d'hébergement en marchand et non-marchand ;
- Excursionnistes (visiteurs sans nuitée hors domicile) ;
- Résidents de la zone (qui peuvent également être usagers du site) ;
- Transit (parfois plus ou moins discernable des autres visiteurs, avec une frontière floue entre un tourisme itinérant et le transit à travers la zone).

Parmi ces catégories, on peut encore distinguer les visiteurs selon les motifs de visite, qui peuvent être :

- Des motifs de tourisme ou de loisirs ;

- Des motifs d'affaire : personnes travaillant dans la zone (ex. personnels des entreprises et services publics présents sur le site), ou se rendant dans la zone pour des motifs professionnels (ex. artisans) ;
- D'autres motifs, par exemple : visites rendues à des résidents du site (amis, famille), courses...

Enfin, on peut aussi catégoriser ces visites plus précisément, selon le type d'activité :

- Randonneurs, promeneurs ;
- Sites payants (musées...) ;
- Activités payantes (visites guidées, activités sportives encadrées, locations...) ;
- Chasse, pêche, cueillette ;
- Autres activités diffuses (visites de villes, points de vue, pique-nique, restaurants, photographie/peinture, diverses activités naturalistes...) ;
- Transit à travers la zone ;
- Visites à des amis, à de la famille ;
- Activités professionnelles.

Ainsi s'agit-il de définir précisément l'objet du suivi. Par exemple, on pourrait ne considérer que les activités de loisirs et de tourisme, en incluant tous les deux types de visiteurs : touristes et excursionnistes, qui passent plus de 2 heures sur le site.

### 1.3. METHODES DE COMPTAGE

Différents modes d'acquisition des données de fréquentation existent, et coexistent le plus souvent dans les systèmes d'information ou observatoires mis en place. Ils se distinguent tout d'abord par le type d'acquisition : direct (comptage des visiteurs, billetteries...), ou indirect via des proxys tels que des nombres de véhicules, des signaux de présence via des registres (hébergements, offices de tourisme...), ou encore via les réseaux sociaux, signaux, dans ce dernier cas, très partiels, qu'il faudra extrapoler pour estimer la fréquentation réelle. Enfin, on peut également distinguer des modes d'acquisition ponctuelle (par ex. de visu sur le terrain) et des modes d'acquisition en continu, par exemple via des dispositifs de comptage automatique (ex. capteurs).

Tableau 1 - Différents modes d'acquisition des données de fréquentation, in (Spenceley et al., 2021)

Direct	<ul style="list-style-type: none"> <li>▪ Counts relying on researchers directly counting.</li> <li>▪ Observations are made at the site, or observing video camera recordings, or using observations from the air.</li> <li>▪ Number of visits or entrants can be counted directly.</li> </ul>
Indirect	<ul style="list-style-type: none"> <li>▪ Inferred counts to provide on-site estimates (e.g. social media posts).</li> <li>▪ Options include counting parking/entrance fees, permits/licences, guest records at accomodation, entries in guestbooks, trail logs, signs of use and social media posts.</li> <li>▪ Numbers of visits or visitors can be counted by permits issued for people entering the site.</li> </ul>
Automatic	<ul style="list-style-type: none"> <li>▪ Counts with mechanical and electronic devices of the numbers of visitors on-site.</li> <li>▪ Counters include traffic counters, and tumstiles of video counters.</li> <li>▪ Entrants can be counted with these devices.</li> </ul>

Tableau 2 - Méthodes de comptage [source : liste de 16 méthodes de comptages sur les espaces littoraux et insulaires terrestres in (Le Corre et al., 2012), légèrement réduite et modifiée] En italique, les sources disponibles dans notre étude

<b>Méthodes de comptages directes</b>
<p><b>Comptages manuels par du personnel sur site</b> Le comptage est réalisé directement sur le site par une équipe de scientifiques ou d'agents à partir de points fixes ou en se déplaçant le long d'un itinéraire prédéterminé. La donnée peut être enregistrée par des compteurs à main ou reportée sur des supports adaptés (fiches d'observation, dictaphones, etc.) avant d'être intégrée à une base de données.</p>
<p><b>Survols avec prises de vues photographiques</b> Lors d'un survol de l'espace étudié et à partir d'un plan de vol prédéfini, un observateur réalise des photographies afin d'identifier et de quantifier les activités humaines. Les clichés sont ensuite photo-interprétés ; les données sont généralement intégrées à un SIG ou à une autre base de données adaptée (Excel, Access, etc.).</p>
<p><b>Traitement des images satellitaires</b> La méthode propose de s'appuyer sur l'interprétation des images satellitaires afin de caractériser et quantifier la fréquentation terrestre.</p>
<p><b>Comptages automatiques à partir de matériels spécialisés (capteurs infra-rouge, dalles acoustiques, etc.)</b> Le comptage est réalisé à partir d'appareils choisis et disposés de façon stratégique sur l'ensemble du site en fonction des usages que l'on souhaite étudier (trafic routier, trafic de vélos, trafic piétons). Les données sont recueillies en continu tout au long de la durée de fonctionnement des appareils.</p>
<p><b>Comptages à partir de photographies automatiques ou de caméras vidéo</b> L'appareil photographique ou la caméra vidéo sont disposés de façon stratégique sur le site. Ils fonctionnent en continu ou se déclenchent lors du passage des visiteurs.</p>
<b>Méthodes de comptages indirectes</b>
<p><b>Evaluation de la fréquentation par indices de présence</b> Méthode d'évaluation de la pression de la fréquentation qui s'appuie sur les indices et les signes laissés par les visiteurs sur le site : traces de passage, piétinement de la flore, érosion des sentiers, dépôt de déchets, etc. L'évaluation est réalisée sur le site à partir d'observations directes ou à partir de photographies aériennes.</p>
<p><b>Traitement des données fiscales telles que la taxe de séjour</b> Elle permet d'extrapoler nombre de visiteurs, nombre de nuitées, taux d'occupation...</p>

**Traitement des données de gestion comptables produites par les gestionnaires**

Il s'agit de s'appuyer sur les données générales produites naturellement par certaines structures-gestionnaires : vente de tickets d'entrées, nombre de sorties organisées, réservations, etc.

**Traitement des données relatives aux permis et aux licences de pratique d'une activité professionnelle, récréative ou touristique**

La pratique de certaines activités professionnelles, récréatives ou touristiques nécessite parfois de disposer d'un permis ou d'une licence (ex : chasse). Dans certains cas, ce sont les aires protégées elles-mêmes qui imposent ces autorisations. Le traitement de ces données permet alors de disposer d'informations générales sur l'évolution temporelle et quantitative d'une activité.

**Traitement des données des téléphones portables**

Sur le principe, les téléphones portables s'apparentent à des transpondeurs. Chaque appareil détient un code unique (ID) et renvoie en continu un signal aux antennes relais les plus proches. En s'appuyant sur ces informations, il est alors possible de quantifier mais également de géolocaliser, les propriétaires de téléphones portables dans un espace considéré.

**Enquêtes par voie postale, électronique ou par internet**

Enquêtes réalisées à petite échelle qui permet d'estimer la proportion de la population d'un pays, d'une région ou d'un département qui fréquentent un territoire ou un milieu spécifique (ex : la forêt, le bassin d'Arcachon) ou qui pratiquent une activité de loisir particulière (ex : la pêche à pied, la plaisance).

**Méthode de la variation de consommation de farine**

La méthode part du constat que le pain est « *d'un usage indispensable, fréquent et constant* ». Aussi, elle propose de s'appuyer sur les consommations de farine afin de mettre en évidence les variations de la fréquentation touristique sur un territoire donnée (une région, un département).

**Inscription volontaire sur les registres des aires protégées**

Traitement des données fournies par les visiteurs eux-mêmes sur les livres d'or, les livres de suggestions disponibles le long du parcours dans les espaces protégés (accueils, gîtes, points d'information, etc.)

Parmi les méthodes les plus récentes permettant un comptage en continu des visiteurs, on peut évoquer les services, développés par des opérateurs de téléphonie mobile, de suivi des visiteurs d'une zone donnée à partir de la détection des téléphones (ex. Orange Flux Vision, en France), ou encore le comptage par caméra (Breyne et al., 2021).

Des approches plus récentes encore, qui se développent actuellement beaucoup, notamment pour l'information qui l'accompagne, réside dans l'exploitation des photos et commentaires déposés sur les réseaux sociaux (Tenkanen et al., 2017), ou encore sur l'exploitation des données GPS, notamment pour étudier la circulation dans les sites (Harder et al., 2008).

Enfin, et surtout, selon le mode d'acquisition des données, l'évaluation de la fréquentation se fera à différents niveaux, chacun présentant des avantages et des inconvénients, et étant plus ou moins adaptés au problème posé : passages, visites, présence journalière sur la zone, cumul de visites, nombre de visiteurs (Tableau X).

Entrant	<ul style="list-style-type: none"> <li>▪ A person going into a protected area for any purpose.</li> <li>▪ The number of entrants includes all recreational visits and other activities (e.g. people just driving through, local people passing through a corner of the park, or the daily activities of workers).</li> <li>▪ Entrant figures can overestimate the recreational use of the park.</li> </ul>
Visits	<ul style="list-style-type: none"> <li>▪ The number of times a person goes into a protected area for recreational purposes.</li> <li>▪ For example, a person who enters a protected area twice in a day would be reported as two visits.</li> </ul>
Visitors	<ul style="list-style-type: none"> <li>▪ A person who visits a protected area for recreational purposes.</li> <li>▪ For example, a person who enters a protected area twice in a day is one visitors.</li> </ul>
Visitor days	<ul style="list-style-type: none"> <li>▪ The total number of days that a visitor stays in a protected area.</li> <li>▪ One overnight stay, or twelve visit hours of a visitor in a protected area is reported as one visitor day.</li> </ul>

Tableau 3 - Différentes mesures de la fréquentation, in (Spenceley et al., 2021)

#### 1.4. PROJET

Le projet VALO-SITES (2019-2022) a pour objectif de développer et tester des méthodes d'évaluation des effets socio-économiques des politiques de gestion et valorisation de sites naturels, en s'appuyant sur les données collectées sur la Chaîne des Puys-Faille de Limagne (CPFL). Il est structuré en plusieurs volets, notamment un volet d'études sectorielles permettant d'alimenter un second volet dont l'objectif est d'établir des estimations économiques (Figure 1).

La tâche T1 du premier volet « Études sectorielles » (encadrée en rouge sur la figure1) est consacrée à l'étude de la fréquentation touristique. Cette tâche est elle-même découpée en deux sous-tâche : la sous-tâche T1A portant sur la quantification des visiteurs, et la sous-tâche T1B portant sur la qualification des visiteurs, à partir d'une large enquête réalisée sur 1011 visiteurs de la Chaîne des Puys, entre mai 2019 et février 2020.

Ce travail correspond à la sous-tâche T1A (encadrée en rouge, sur fond rouge, dans la figure 1), dont l'objet est l'étude des méthodes de comptage des visiteurs mobilisables dans des sites naturels, en s'appuyant sur le cas d'étude de la CPFL, avec l'objectif conjoint de fournir une estimation de la fréquentation totale du site sur l'année T0 (2019).

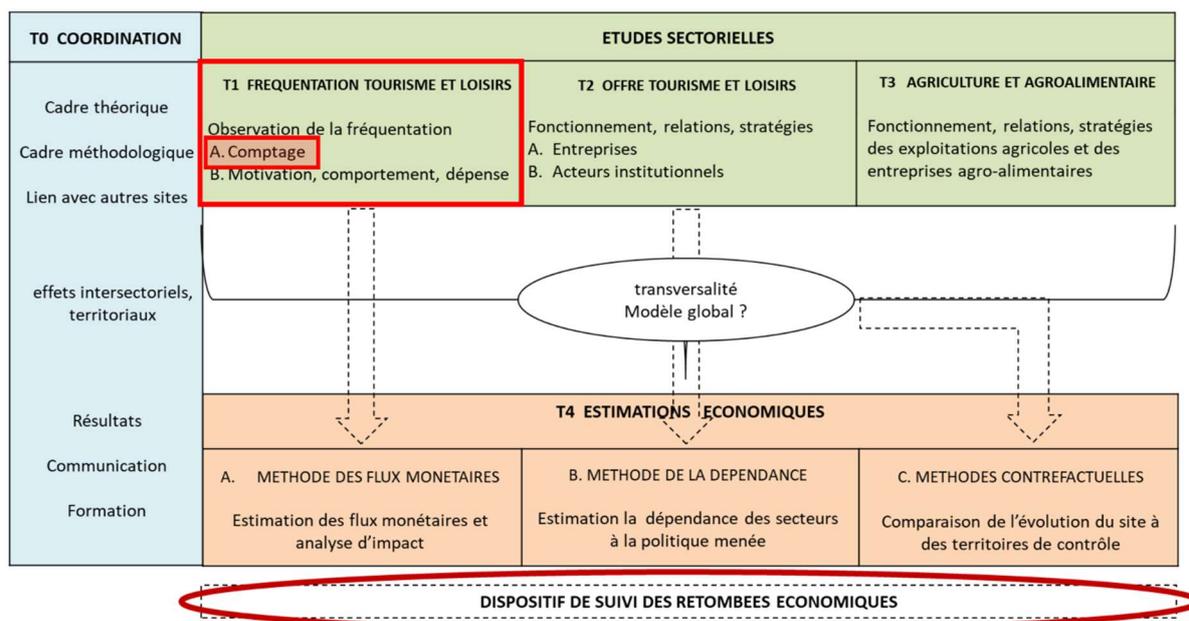
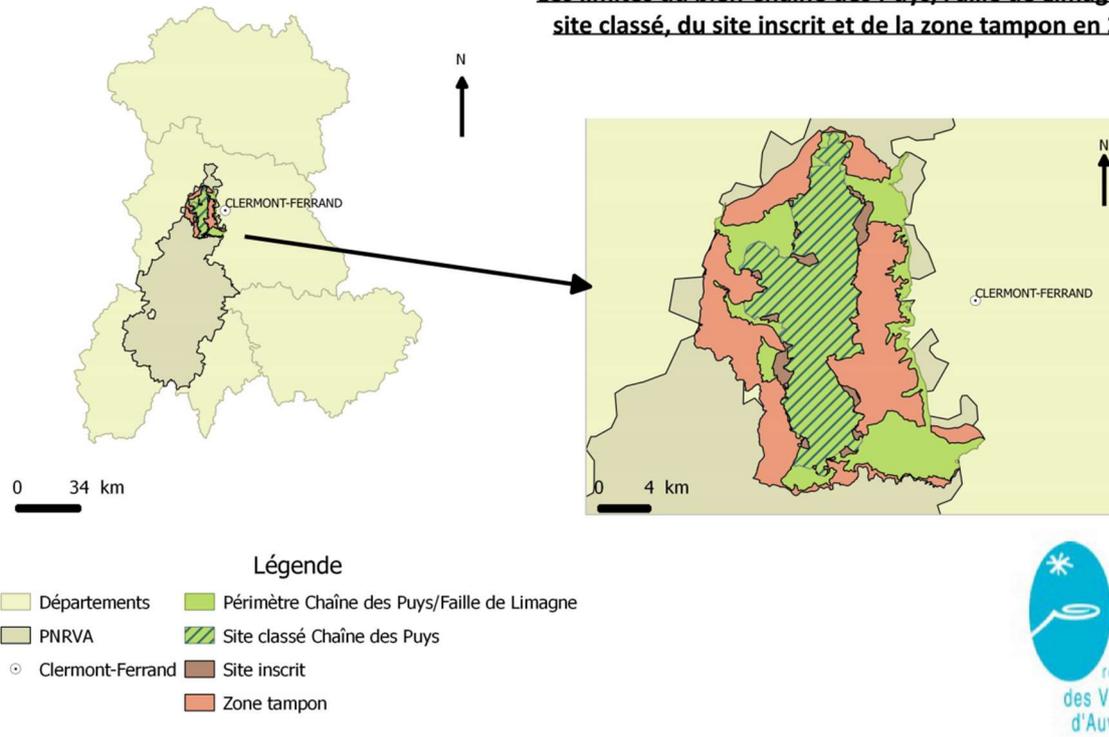


Figure 1 - Schéma général du projet VALO-SITES

## 1.5. ZONE D'ETUDE

La zone d'étude correspond au site Chaîne des Puys – Faille de Limagne, inscrite le 2 juillet 2018 au patrimoine mondial de l'UNESCO, dans la liste des sites naturels, en tant que haut lieu tectonique, en incluant la zone tampon, définie à des fins de préservation du site. Elle inclut la zone classée en 2000 au titre de la loi de 1930, et la zone labélisée Grand Site de France, étendue lors de son renouvellement en 2021. Enfin, elle fait partie du Parc Naturel Régional des Volcans d'Auvergne (Figure 2).

**Les limites du bien Chaîne des Puys/Faille de Limagne, du site classé, du site inscrit et de la zone tampon en 2015**



Source : Géofla Conception : Amandine MAUREAU

Figure 2 - Carte de localisation du site Chaîne des Puys - Faille de Limagne, (source : (Moreau, 2015))

La zone inclut les différentes formations d'intérêt du bien inscrit à l'UNESCO : la Chaîne des Puys, la Faille de Limagne, la montagne de la Serre et le plateau des Dômes (Figure 3).

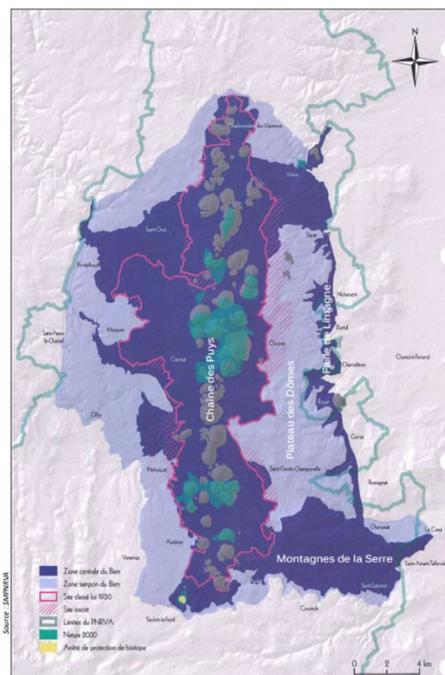


Figure 3 - Carte des différents éléments d'intérêt du bien inscrit, et périmètres des zones classées et protégées (source : SMPNRVA, in (Moreau, 2015))

Il s'agit d'une **zone habitée**, recoupant plusieurs dizaines de communes (Figure 4), et comptant plus de 25 000 résidents, avec des communes résidentielles du péri-urbain de Clermont-Ferrand, telles que Saint-Genès-Champanelle, Aydat, Orcines, se traduisant notamment par des travailleurs, des personnes rendant visite aux résidents, etc., qui doivent être distingués des visiteurs de tourisme et de loisirs. De plus, ces communes disposent d'hébergements touristiques, impliquant la présence de touristes logeant dans les communes du site, pouvant passer leurs journées, dans ou en dehors du site.

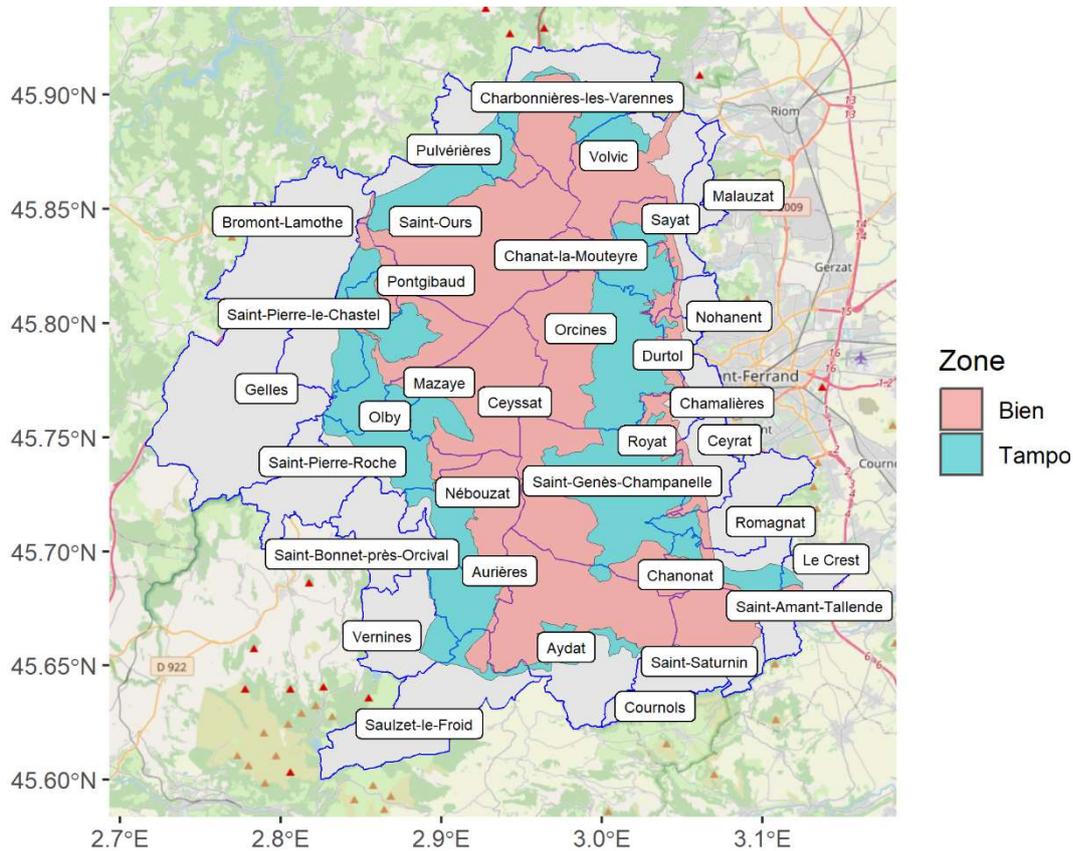


Figure 4 - Carte des communes, incluses ou jouxtant la zone d'étude

Une autre particularité réside dans l'ouverture du site, qui est un lieu d'attractivité résidentielle, car proche de Clermont-Ferrand, mais également un lieu de transit, en raison de la présence de nombreuses routes traversantes (Figure 5), dont certaines très fréquentées, en particulier la RN89 desservant les zones résidentielles à l'Ouest de Clermont-Ferrand (Saint-Genès-Champanelle, Aydat, Rochefort-Montagne) et permettant également de joindre les autoroutes A75 (Montpellier) et A71 (Lyon) à l'autoroute de Bordeaux, en évitant le détour par le nord de Riom (Cf. Figure 5 : plus de 7 000 par jour au col de la Ventouse, et plus de 11 000 par jour entre Theix et Clermont-Ferrand), et la D942 desservant les zones résidentielles au Nord-Ouest de Clermont-Ferrand (Orcines, Ceysnat) (Cf. Figure 5 : environ 7 000 par jour au niveau du parking des Goules). La présence de nombreux résidents, les mobilités importantes, notamment avec Clermont-Ferrand, font de cette zone une entité très ouverte, à la différence de sites naturels plus enclavés ou moins habités. Par conséquent, beaucoup de visiteurs peuvent être soit uniquement en transit, soit être venus pour des raisons autres que des motifs de tourisme et de loisirs, ce qui complexifie également la tâche d'estimation de la fréquentation.



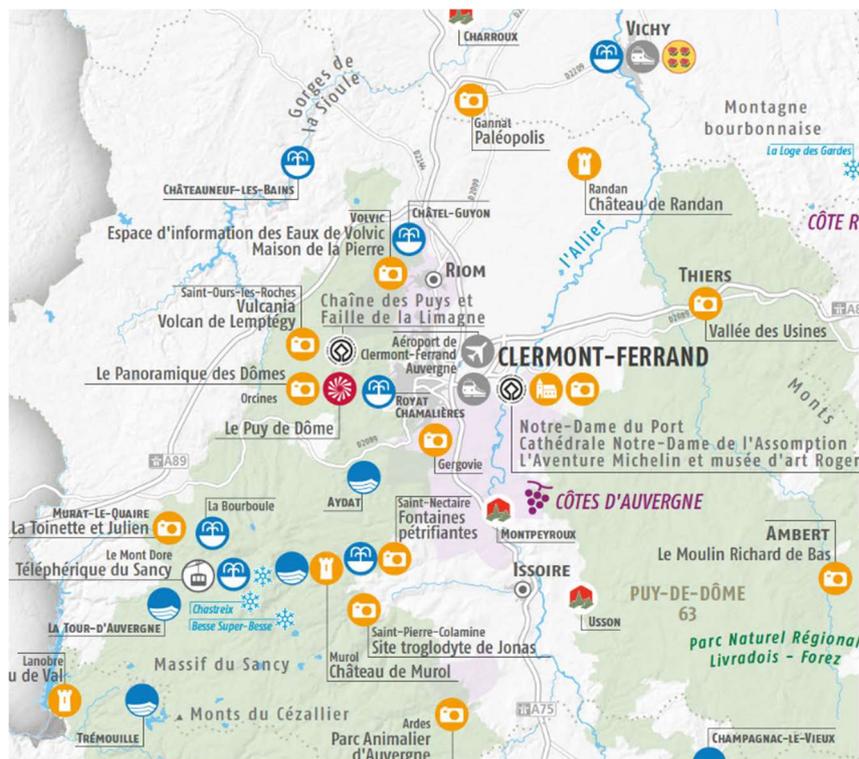


Figure 6 - Extrait de la carte touristique Auvergne-Rhône-Alpes (source : Auvergne-Rhône-Alpes Tourisme)

La zone d'étude dispose également d'un réseau complexe de sentiers : de nombreux sentiers officiels, parfois très aménagés, d'autres beaucoup moins, ainsi que de nombreux sentiers et chemins très utilisés par les promeneurs, en particulier par les locaux. Ces différents sentiers peuvent être très interconnectés, permettant de nombreuses variantes de circuits, ce qui rend le comptage exhaustif des visiteurs, par des méthodes automatique (capteurs, caméras) quasi-impossible. La zone d'étude est, depuis plusieurs années, très bien équipée en éco-compteurs, mais de nombreux sites ne le sont pas encore (par ex. montagne de la Serre, Aydat), ainsi que de nombreux circuits, notamment ceux ne montant pas aux sommets des puys (par ex. le tour du Puy de Dôme, très fréquenté). En outre, certains sites ne se prêtent sans doute pas à ce mode de comptage (par ex. la plage d'Aydat).

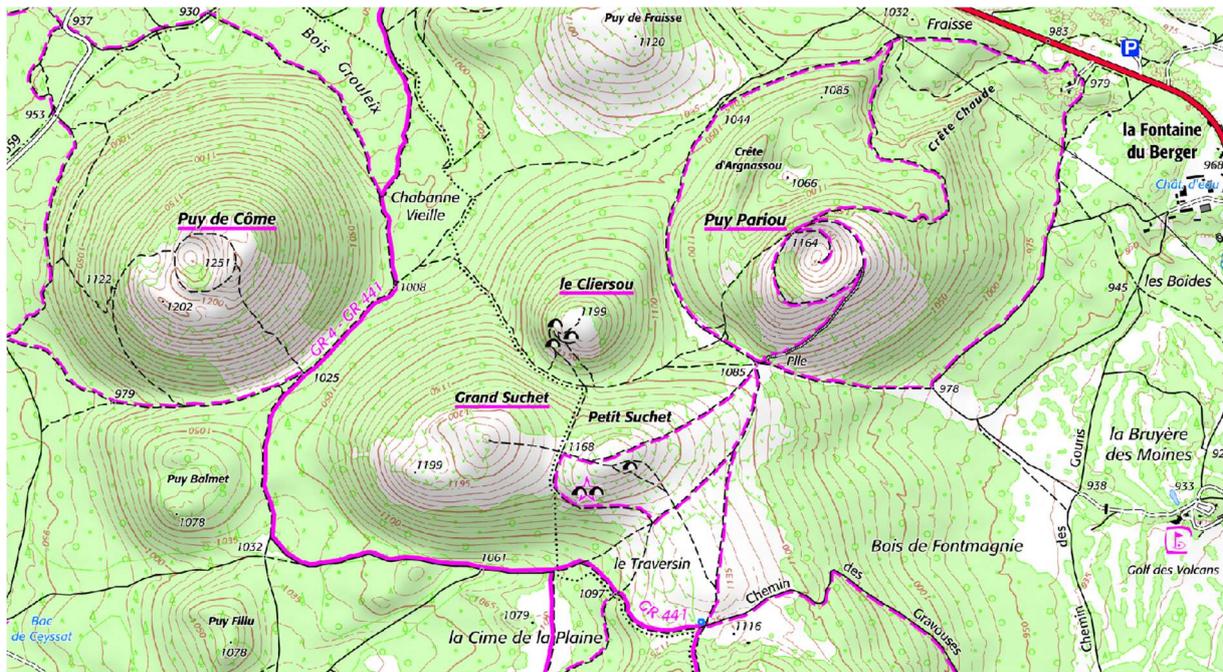


Figure 7 - Extrait de la carte scan 25 centrée sur la zone du Puy de Pariou (source : IGN)

Une autre particularité de la zone d'étude provient du fait que **toutes les composantes de fréquentation** évoquées dans la partie précédente sont ici présentes :

- Randonneurs (essentiellement sur et autour des puys) ;
- Promeneurs, joggers et VTTistes (puys, mais surtout sentiers bas autour des puys, tour du lac d'Aydat...)
- Autres loisirs et activités sportives (balades à cheval, nautisme à Aydat, parcours sportifs, parapente, montgolfière...)
- Plage (Aydat) ;
- Chasse, pêche (La Cassière) ;
- Panoramique des Dômes, combiné - ou pas - avec la montée ou la descente par les sentiers, et utilisé par les parapentistes ;
- Sites payants (Vulcania, Volcan de Lemptégy, Maison de la Pierre à Volvic) ;
- Sites d'information (Maison du Parc à Montlosier, Espace d'information à Volvic, Office du Tourisme d'Aydat) ;
- Activités diffuses : pique-nique, restaurant, visite de ville, stationnement de camping-car, circuits en voiture vers des points de vue, renseignement auprès des lieux d'information ou des sites payants...

Or, certaines de ces composantes seront probablement plus difficiles à quantifier. De plus, la fréquentation du site est globalement très mélangée, à la fois très touristique et très fréquentée par les résidents du département, notamment du fait de la proximité de la métropole de Clermont-Ferrand, et peut-être des départements voisins (excursions).

## 1.6. OBJECTIFS

L'objectif principal de la tâche T1A du projet Valo-Sites est d'étudier les méthodes de suivi de la fréquentation sur des sites naturels. A partir du cas d'étude de la Chaîne des Puys – Faille de Limagne (CDPFL), il s'agit donc de s'interroger sur les approches envisageables pour réaliser ce suivi, afin d'obtenir des estimations les plus exhaustives possibles, de façon reproductible.

En réponse à une attente du conseil départemental du Puy de Dôme, un second objectif est également de proposer, pour l'année 2019, qui pourrait être considérée comme l'année 0 d'un futur suivi, une méthode d'estimation de la fréquentation annuelle globale sur CDPFL, à partir des données actuellement disponibles (éco-compteurs, billetteries).

En premier lieu, il sera nécessaire de voir dans quelles mesures il est possible de réaliser cette estimation à partir des données déjà collectées sur le site. Dans un second temps, d'autres approches seront envisagées, et autant que possible, testées et comparées entre elles.

Le **premier objectif** consiste à produire une première estimation de la fréquentation totale du site, pour l'année 2019 (année zéro du suivi), à partir des données actuellement disponibles (éco-compteurs, billetteries), afin d'avoir une première idée du nombre annuel de visites, nécessaire notamment à l'évaluation des retombées économiques, mais également d'obtenir un premier élément de comparaison, pour évaluer d'autres méthodes d'estimation. En parallèle, l'exercice doit permettre d'identifier les éventuelles difficultés et limites de cette approche.

Pour ce faire, il s'agit, en premier lieu, de recenser l'ensemble des données disponibles, et d'étudier la pertinence, la représentativité et l'exhaustivité de ces données, c'est-à-dire de les confronter à l'ensemble des données de ce type, qui seraient nécessaires pour réaliser parfaitement l'exercice.

Ensuite, il faudra évaluer la qualité de ces données, en analysant chaque source, pour identifier les données manquantes et les anomalies (données aberrants, incohérences...).

A l'issue de cette étape, il sera éventuellement nécessaire de réaliser des corrections, voire des imputations et des extrapolations, afin de disposer de données complètes permettant une estimation de la fréquentation globale sur une année complète.

Idéalement, il faudra essayer d'exprimer la fréquentation, à partir d'un nombre de visites sur le site, en un nombre de visiteurs, puis essayer d'aboutir à une première qualification de cette fréquentation, notamment en distinguant touristes et excursionnistes.

Enfin, la production de données de fréquentation à une résolution temporelle plus fine pourrait permettre de mieux suivre la fréquentation dans le temps, d'étudier plus finement sa saisonnalité, mais également de mieux calibrer les enquêtes de terrain, voire de modéliser la fréquentation, pour prédire à court-terme ou essayer d'expliquer des phénomènes précis.

Parallèlement, un **second objectif** consiste à étudier d'autres approches envisageables pour le suivi de la fréquentation. En 2019, deux options ont été proposées pour tester des

approches alternatives : une expérimentation via les comptages routiers, et un test du service Flux Vision d'Orange.

Contraints par le budget alloué à cette tâche du projet, ces deux tests se sont limités, pour les comptages routiers, à une expérimentation d'environ un mois réalisée gratuitement, durant l'été 2019, par la Direction des Routes du département du Puy de Dôme et, pour Flux Vision, par l'achat de trois mois de données, couvrant la période d'août à octobre 2019.

## 1.7. PLAN

Dans une première partie, nous présenterons les différentes sources de données disponibles, les traitements effectués, et enfin l'estimation globale de la fréquentation sur CDPFL en 2019, à partir des visites sur les sites dénombrables.

Ensuite, dans la seconde partie, nous décrirons les approches alternatives qui ont été testées : les comptages routiers, et le service Flux Vision d'Orange.

Puis, dans une troisième partie, nous présenteront les comparaisons entre ces estimations, c'est-à-dire entre l'estimation globale à partir des sites dénombrables et celle obtenue avec la méthode Flux Vision. Les autres informations sur les visiteurs fournies par Flux Vision seront confrontées avec celles collectées lors de la grande enquête Valo-Sites.

Enfin, pour conclure, nous synthétiserons les principaux enseignements de ces analyses, et nous évoquerons quelques propositions et pistes pour aller plus loin dans la mise en place d'un dispositif de suivi de la fréquentation.

## 2. ESTIMATION A PARTIR DES SITES « DENOMBRABLES »

### 2.1. DONNEES DISPONIBLES

La connaissance générale de la fréquentation sur la zone CDPFL, en considérant le Bien inscrit à l'Unesco et la zone tampon, permet d'identifier un ensemble de sites que l'on qualifiera de « dénombrables », dans la mesure où l'on pourrait potentiellement quantifier le nombre de visiteurs spécifiques, à partir de données disponibles (agrégation des données d'éco-compteurs, cumuls d'entrées annuels sur les sites payants).

Les principales attractions identifiées peuvent constituer la base d'un dénombrement de la fréquentation globale de la zone CDPFL. Ainsi pourrait-on reconstituer la fréquentation globale à partir de la fréquentation des principales composantes : circuits de randonnée, Panoramique des Dômes, sites payants (Vulcania, le Volcan de Lemptégy, la Maison de la Pierre à Volvic), et plage d'Aydat, qui représentent les principales attractions de la CDPFL.

La question est de savoir si les données disponibles, via les éco-compteurs et les billetteries, sont suffisantes pour estimer la fréquentation globale. Reste qu'une partie de la fréquentation ne sera jamais comptabilisée : sites non équipés d'éco-compteur ou de billetterie, sentiers alternatifs non-équipés d'éco-compteur, fréquentation diffuse, hors attractions.

Enfin, une simple agrégation des visites par site dénombrable présente un risque de doubles comptes, en particulier lorsqu'un même visiteur fréquente plusieurs sites de la zone dans une même journée.

Le recensement des données disponibles sur ces sites dénombrables fait apparaître, pour l'année 2019, les sources suivantes (tableau X) :

- 10 sites de randonnées sur les puys équipés par 16 éco-compteurs, avec des données infra-horaires, dans les deux sens de passage ;
- Des données par train (infra-horaire) pour le Panoramique des Dômes (avec des données mensuelles ventilées par origine : pays/département) ;
- Uniquement le nombre total d'entrées sur l'année pour les sites payants, à l'exception de Lemptégy, où l'on dispose également des données mensuelles ;
- Enfin, aucune données de fréquentation pour le Lac d'Aydat, à l'exception des données journalières d'entrée dans l'office de tourisme d'Aydat.

Tableau 4 - Données disponibles pour les sites dénombrables

Source	Type	Grain	Origines	Complément
<b>Eco-compteurs</b>	passages (IN + OUT) 16 EC / 10 sites	infra-horaire		estimation pertes (PNR)
<b>Panoramique</b>	passages (A/R)	infra-horaire (par train)	par mois	
<b>Lemptégy</b>	entrées	mensuel	sur l'année	entrées et origines par jour sur le mois d'août 2019
<b>Vulcania</b>	entrées	annuel		8 dates en août 2019
<b>Maison de la Pierre (Volvic)</b>	entrées	annuel		
<b>Espace Information (Volvic)</b>	entrées			
<b>Aydat</b>				entrées OT (+ origines) estim. pic estival instantané
<b>Maison du PNR</b>				Entrées estimées

### 2.1.1. ÉCO-COMPTEURS

#### FONCTIONNEMENT

Les éco-compteurs sont des systèmes de comptage automatique des flux de piétons et de cyclistes. Plusieurs technologies existent : des dalles sensibles aux variations de pression dissimulées au sol, ou des capteurs infra-rouge détectant le passage des personnes.

Sur la CDPFL, il s'agit de capteurs infra-rouge (Figure 8) dissimulés en bordure de chemin, le plus souvent dans des troncs (Figure 9). Ils permettant de détecter les personnes et leur sens de circulation. Les données sont directement accessible sur une interface en ligne, Eco-Visio (Figure 10). Le coût est d'environ 25 000 euros pour 10 éco-compteurs (Moreau, 2015).

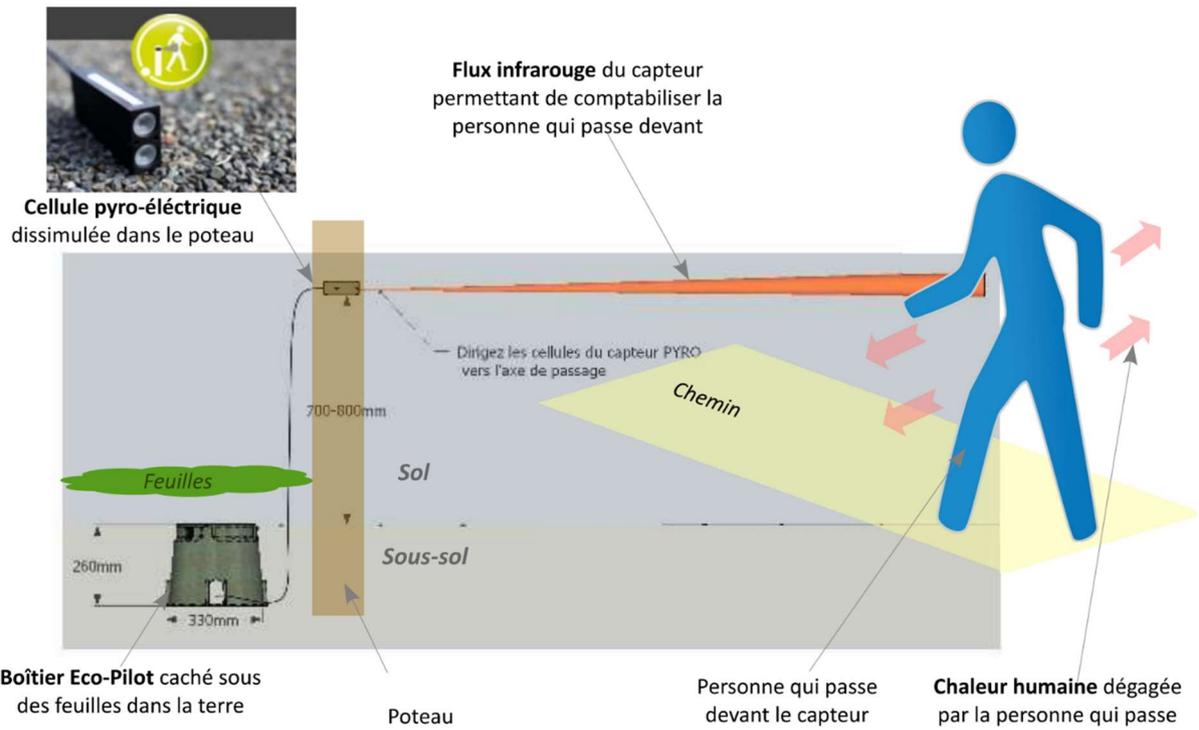


Figure 8 - Schéma du fonctionnement de l'éco-compteur - source : (Moreau, 2015)

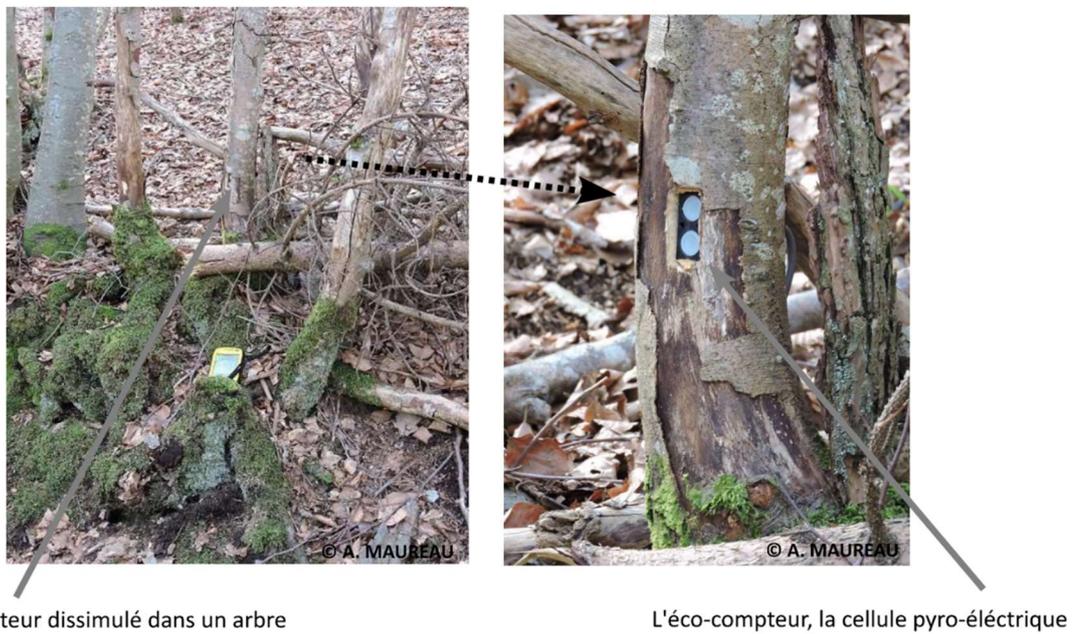


Figure 9 - Photos d'éco-compteur in situ (source : (Moreau, 2015))

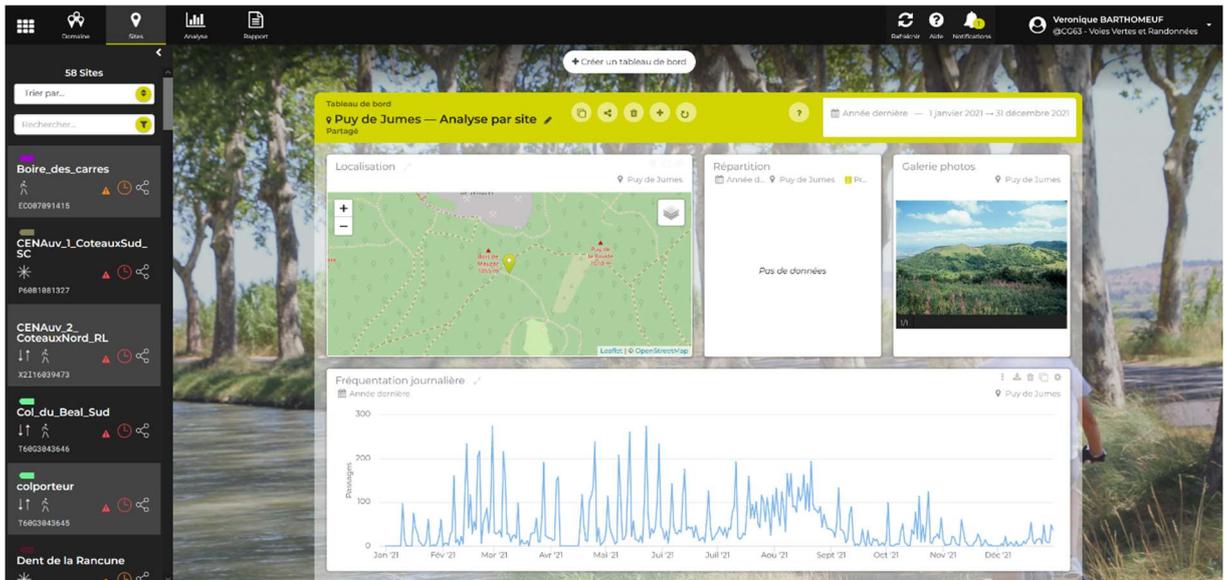


Figure 10 - Copie d'écran de l'interface ecovisio permettant la visualisation et l'extraction des données d'éco-compteur

Parmi les problèmes rencontrés, on peut citer les cas de vandalisme, la dégradation naturelle (météo, câbles rongés par des animaux...), le dysfonctionnement du capteur. Ces cas sont relativement fréquents, et font perdre, chaque année ou presque, plusieurs jours ou semaines de données sur des éco-compteurs de la CDPFL.

En outre, il faut également évoquer le problème du passage de plusieurs personnes simultanément devant le capteur, même s'il est censé, d'après le fournisseur, ne pas affecter pas la mesure (Figure 11). Cela devrait être testé, pour s'assurer de la fiabilité des mesures.

# Capteur PYRO

Compteur de personnes pour comptage en extérieur

## Avantages :

- Détection du sens de passage
- Technologie non-intrusive
- Autonome en énergie (10 ans)
- Étanche (IP 68)
- Portée jusqu'à 15 m
- Facile à camoufler



## Fonctionnement

Le capteur **PYRO** utilise la combinaison d'une **technologie pyroélectrique infrarouge passive** et d'une **lentille haute précision** pour détecter la **chaleur émise par le corps humain**. Grâce à son **extrême sensibilité** et à la **puissance de son algorithme ORION**, le capteur PYRO est capable de **différencier deux personnes se suivant de très près**, et de déterminer leur **sens de passage**. Capteur multifonction, il peut compter **randonneurs, cyclistes, cavaliers, kayakistes, skieurs...** Livré avec ou sans habillage, il suffit d'un mur de pierres sèches ou simplement d'utiliser la végétation pour le **dissimuler**.

Figure 11 - Description du capteur PYRO (site [eco-compteur.com](http://eco-compteur.com))

Le positionnement des éco-compteurs est crucial, notamment pour éviter les fuites, par exemple celles liées à la circulation hors-sentier ; c'est notamment le cas avec l'apparition d'un faux sentier, qui peut être due à plusieurs premiers passages hors-sentier, ou encore à la circulation du bétail. Autre exemple, certains éco-compteurs ayant été mis en place dans le cratère (ex. Puy des Gouttes), ils ne détecteront donc pas les personnes qui restent sur les crêtes. Enfin, lorsque de nombreux chemins alternatifs coexistent, le problème reste insoluble.

Dans une note en ligne sur les systèmes de comptage automatique des personnes (SVIOM Grand Site Gorges du Tarn, de la Jonte, et Causses), sont listées les recommandations suivantes :

« - Choisir des passages étroits et obligés (où les personnes peuvent passer en file indienne), et bien mesurer la largeur du passage avant de passer la commande d'éco-compteur ;

- Privilégier les zones de trafic franc : éviter les zones d'arrêt type point de vue, et les zones trop près des stationnements propices aux allers-retours ;

- Besoin de relever les compteurs une fois par mois, pour vérifier qu'il n'y a pas de problèmes et ne pas perdre des mois de données ;

- Il peut être utile de réaliser des contre-comptages manuels ponctuels (observation visuelle), pour corroborer les données automatiques ;

- Phase test les deux premières années d'installation pour avoir des données fiables exploitables (comparaison sur deux années permettant de vérifier la fiabilité du comptage).

Espaces naturels qui ont fait des études de fréquentation : Puy Mary, PNR Verdon, PNC Vanoise, Ecrins et Pyrénées, CG Ardèche, Gorges du Tarn et de la Jonte, Baie de Somme... » (SIVOM Grand Site Gorges du Tarn, de la Jonte, et Causses, r, 2017)

D'autres auteurs proposent diverses recommandations pour leur fonctionnement optimal, notamment en terme de calibration (Rupf et al., 2008).

---

## DONNEES

Sur la zone d'étude, une vingtaine d'éco-compteurs a été mise en place, dont 16 sont réellement exploitables pour notre étude, portant sur l'année 2019. Hormis l'éco-compteur du sentier des mulotiers, géré par le conseil départemental, l'ensemble des autres éco-compteurs de la zone est géré par le PNR des Volcans d'Auvergne.

Globalement, on peut distinguer 3 zones principales (Figure 12) :

- Les puys du nord de la chaîne (Nugère, Louchadière, Jumes, Coquille, Gouttes, Goules-Sarcoui) ;
- Les puys de la zone centrale (Pariou, Côme, Cliersou, Puy de Dôme) ;
- Les puys du sud de la chaîne (Vache, Lassolas, Vichatel, Combegrasse).

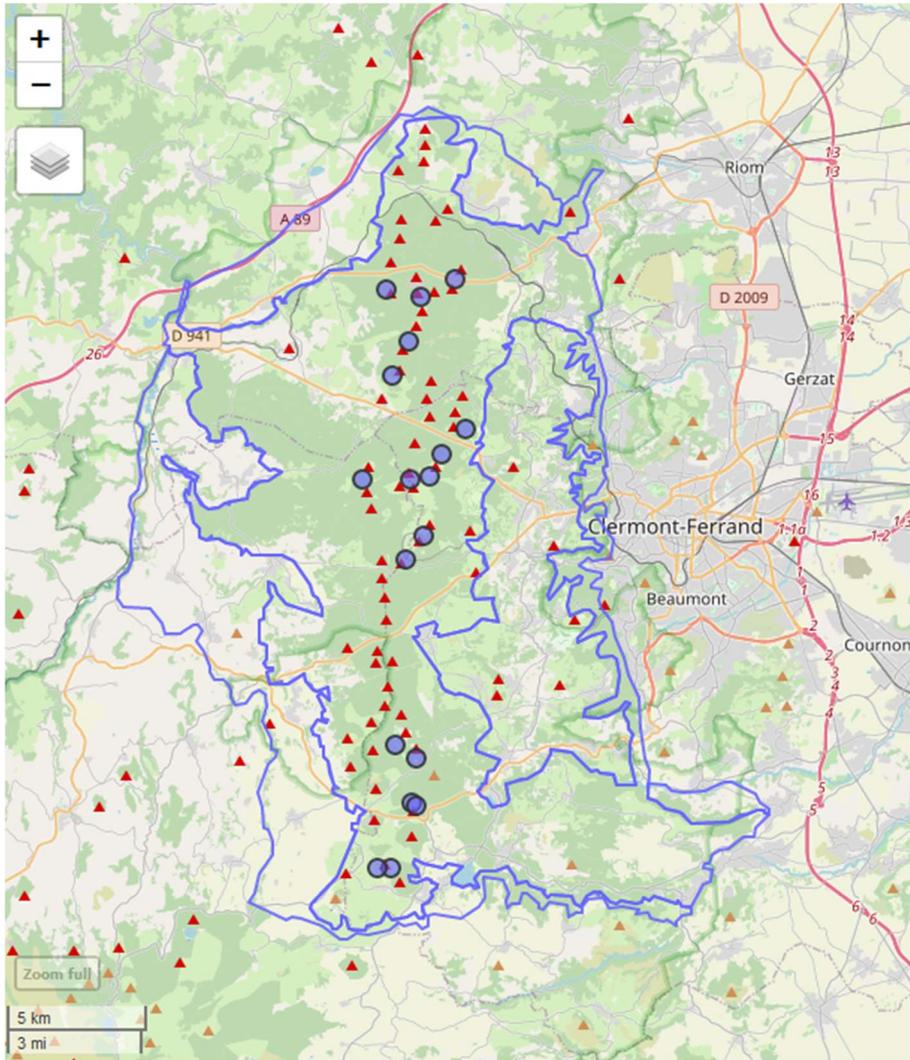


Figure 12 - Répartition spatiale des éco-compteurs sur la zone d'étude (données PNRVA)

On dispose donc de deux séries pour chaque éco-compteur, une par sens de passage, IN et OUT (Figure 13). Sur certains sites, plusieurs éco-compteurs capturent le cheminement des visiteurs. Ainsi, sur Combegrasse, les deux éco-compteurs (Combegrasse-Est et Combegrasse-SO) fonctionnent ensemble ; on voit que les montées par l'un correspondent aux descentes par l'autre. Sur d'autres sites, un seul éco-compteur capture les montées et les descentes, et l'on peut constater une superposition des deux courbes, IN et OUT ; c'est notamment le cas pour Goules-Sarcoui et le sentier des muletiers. Sur d'autres sites, en revanche, soit il n'y a pas de correspondance nette, en raison de circuits plus complexes ou de pertes (sentiers alternatifs), soit le visiteur n'est détecté que dans un sens.

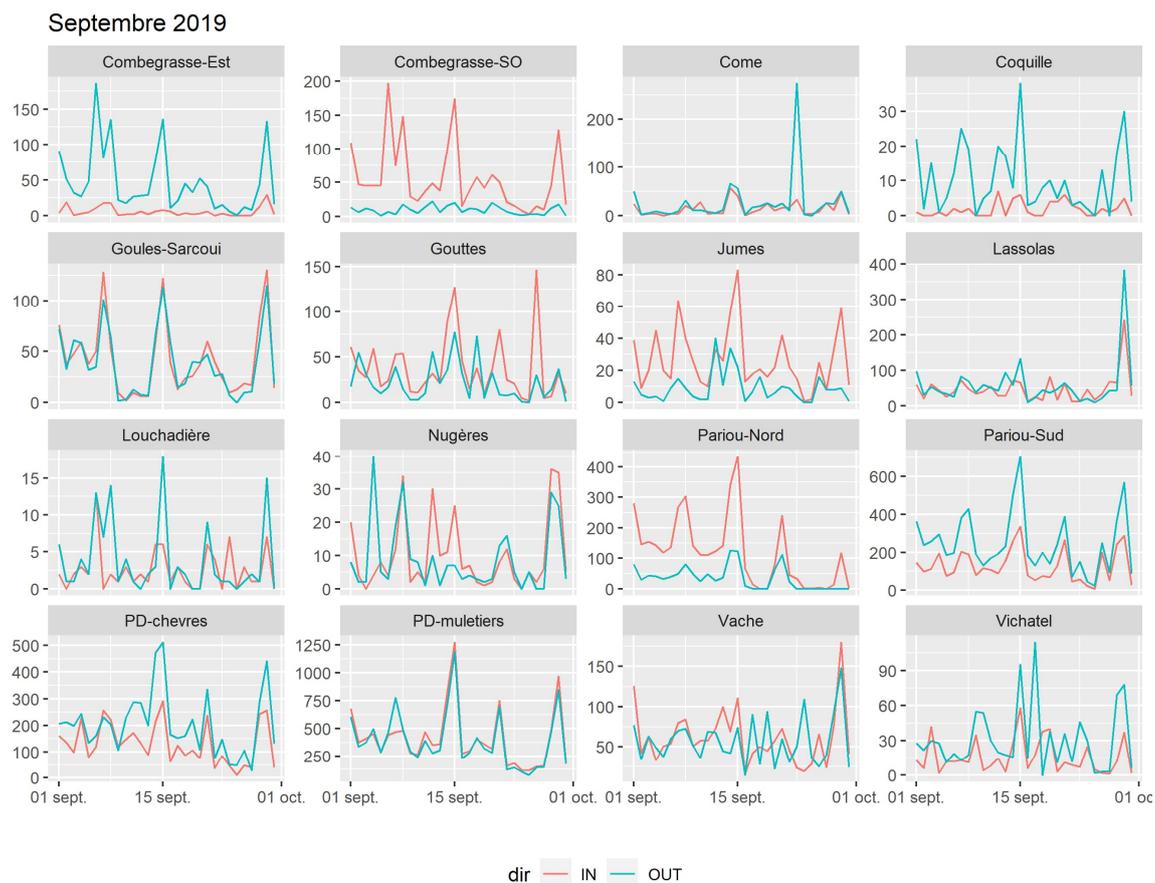
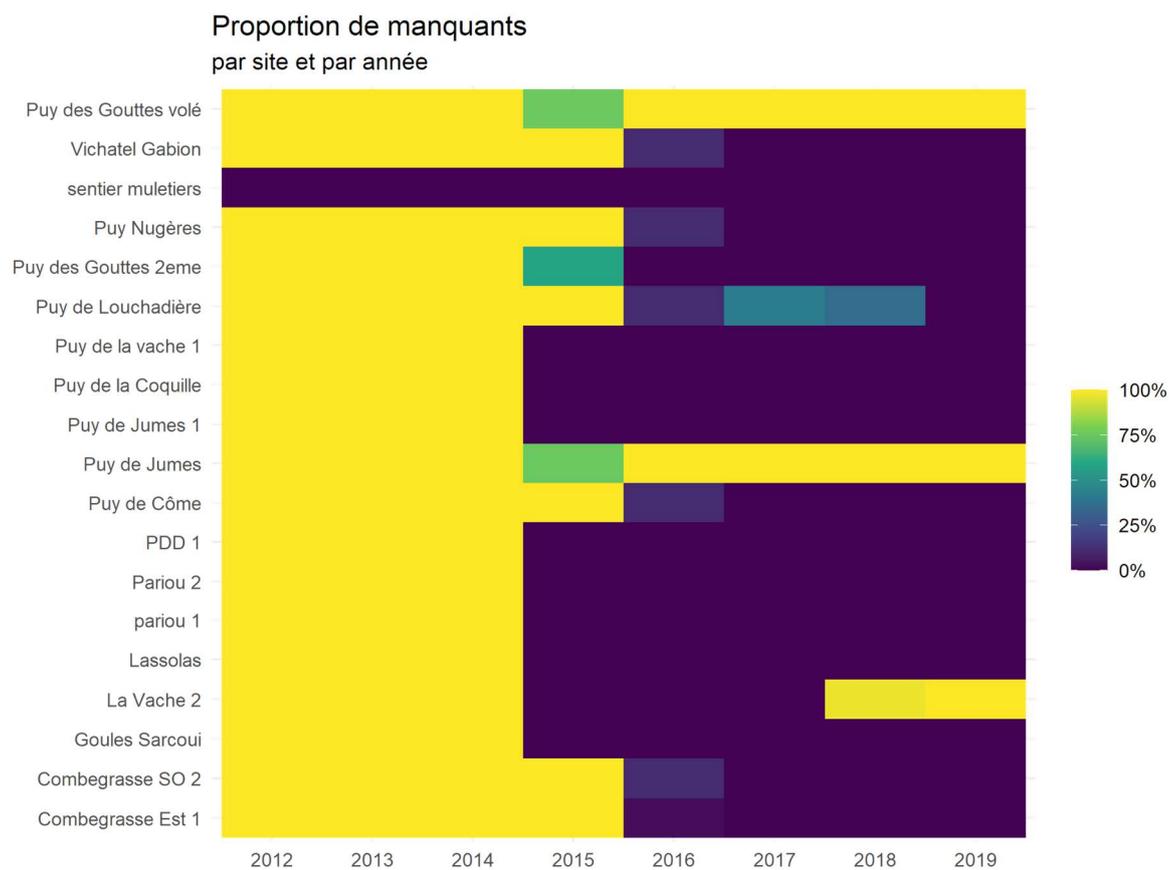


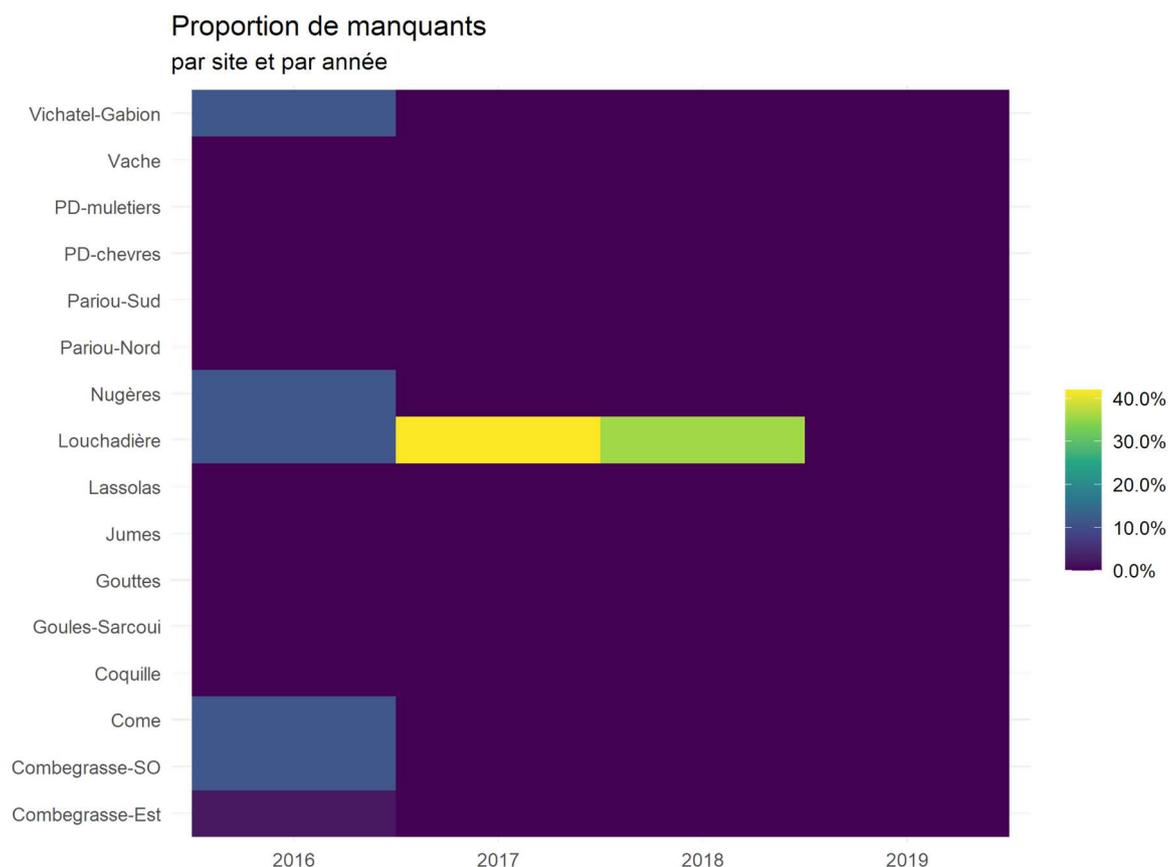
Figure 13 - Séries journalières par sens de passage, sur le mois de septembre 2019, pour les 16 éco-compteurs disponibles

Sur une extraction des données depuis 2012 (Figure 14), on constate que l'EC du sentier des muletiers présente la plus longue série de données. Les autres éco-compteurs ont été mis en place entre 2015 et 2016 (à l'exception de l'EC du Puy de Clersou, mis en place plus récemment). Enfin, certains ont été volé (ex. EC Puy des Gouttes) ou déplacé suite à la fermeture du sentier (ex. EC Vache 2).



**Figure 14 - Proportion de données manquantes par éco-compteurs et par année**

On peut constater que sur ces dernières années on dispose d'une série quasi complète pour une quinzaine d'éco-compteurs, et même 16 sur l'année 2019 (Figure 15).



**Figure 15 - Proportion de données manquantes par année pour 16 éco-compteurs de la zone d'étude (2016-2019)**

## REPRESENTATIVITE

Pour mieux comprendre le contexte, le fonctionnement et l'acquisition des données d'éco-compteurs, nous avons rencontré les gardes nature du PNR au début de l'année 2019.

En février 2021, nous avons eu un nouvel échange (en visio) avec Antony Porte et Benoît Tollet, principalement pour discuter des pertes potentielles dans le comptage des visiteurs sur les différents puys. Sur la base des cartes scan25 (Figure 16 et Figure 17), sur lesquelles sont représentés les principaux chemins, nous leur avons demandé, pour chaque puy, quelle était selon eux la part de randonneur qui pouvait ne pas passer devant l'éco-compteur. Bien qu'ils n'aient pas les éléments pour faire une telle estimation, ils ont accepté de nous proposer un chiffre, à dire d'expert, qui puisse refléter les écarts possibles entre le nombre de passages et le nombre réel de visiteurs, sur la base de ce qu'ils connaissent de la fréquentation de ces puys et les éventuels sentiers alternatifs. Ces chiffres, et les explications données, sont indiqués dans Tableau 5.

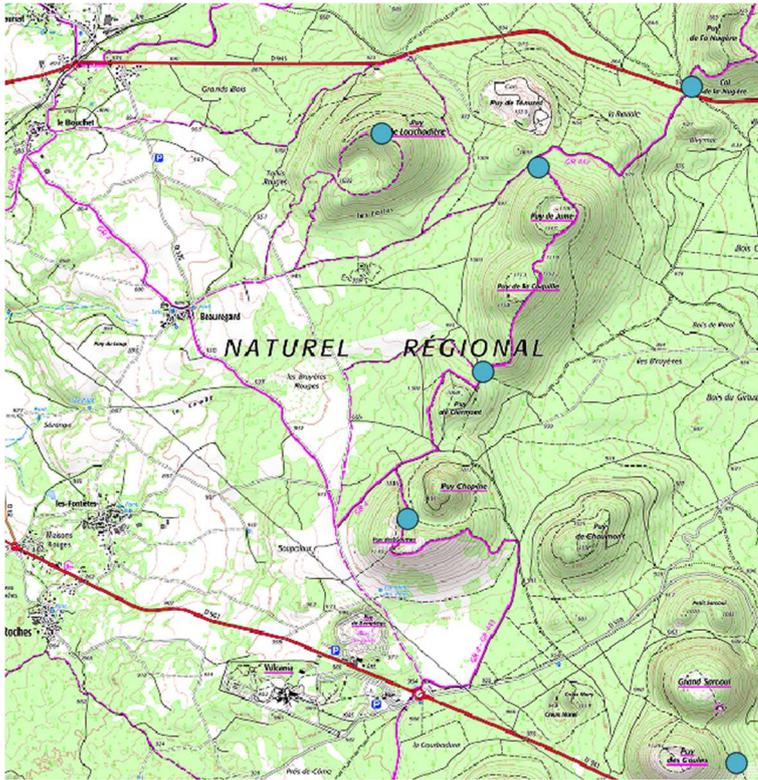


Figure 16 - Carte scan25 centrée sur les puy du nord de la zone d'étude

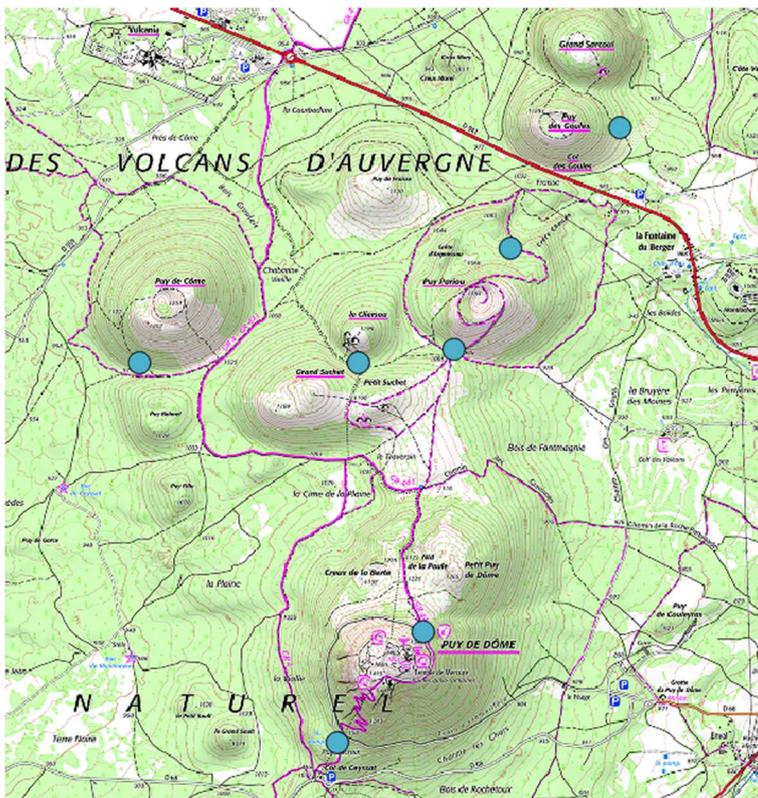


Figure 17 - Carte scan25 centrée sur la zone Pariou - Puy de Dôme

Tableau 5 - Coefficient multiplicatif permettant d'extrapoler le nombre réel de visiteurs sur ou à proximité du puy

Écocompteur	Coefficient	Explication
<b>Puy de Dôme</b>	1.4	Autres circuits (ex. Tour du Puy de Dôme)
<b>Pariou</b>	2	Beaucoup de circuits en bas (ex. Chemin des Gouris)
<b>Côme</b>	2	EC avec le + de pertes (majorité passe une seule fois, voire 0)
<b>Gouttes</b>	2	EC dans le cratère (tout le monde ne descend pas)
<b>Goules-Sarcoui</b>	1.2	Quelques pertes
<b>Jumes-Coquille</b>	1	OK
<b>Vache-Lassolas</b>	1.2	Avec Vichatel
<b>Combegrasse</b>	1	OK
<b>Nugère</b>	2	Uniquement ceux qui viennent d'un côté (50% ?)
<b>Louchadière</b>	3	Beaucoup de pistes non balisées ! (x2, x3 ?)

Lors de ces discussions, nous avons également évoqué les autres problèmes : le vandalisme, le positionnement des éco-compteurs (seulement sur les accès aux puys seulement, et parfois sur la descente dans le cratère (ex. Gouttes), les sentiers non équipés, les passages simultanés et la fiabilité des éco-compteurs, notamment en lien avec le passage simultané (les éco-compteurs les plus concernés nous ont été indiqués, pour d'éventuels tests).

---

### 2.1.2. PANORAMIQUE DES DOMES

Les données du Panoramique des Dômes sont disponibles depuis la mise en service en juillet 2012 sous la forme d'un tableau par jour, mentionnant les nombres de voyageur dans chaque train, à la montée et à la descente, accompagné de l'horaire et du retard éventuel (Figure 18).

	A	B	C	D	E	F	G	H	I	J
1	<b>REGULATION DES TRAINS ET DES VOYAGEURS</b>									
2	<b>31</b>	<b>juillet</b>	<b>2019</b>							
3	<b>METEO</b>									
4	<b>GDS</b>					<b>MDS</b>				
5			T°	Humidité	Vent			T°	Humidité	Vent
6	<i>Matin</i>	Nuageux	HS	HS	30,0Km/h	<i>Matin</i>	Nuageux	10,2°C	87,0%	0,0Km/h
7	<i>Après midi</i>	Nuageux	HS	HS	20,0Km/h	<i>Après midi</i>	Nuageux	14,2°C	61,0%	2,0Km/h
8										
9										
10	<b>Trains départ MDS</b>					<b>Trains départ GDS</b>				
11	Horaires		Retard	Nombre Voyageurs	Cumul	Horaires		Retard	Nombre Voyageurs	Cumul
12	09:00									
13		09:20				09:20				
14	09:40						09:40			
15		10:00				10:00				
16	10:20						10:20			
17		10:40				10:40				
18	11:00						11:00			
19		11:20				11:20				
20	11:40						11:40			
21		12:00				12:00				

Figure 18 - Format des données du Panoramique des Dômes (capture d'écran)

Ces données sont accompagnées de données mensuelles sur l'origine des visiteurs, par pays et par département.

### 2.1.3. SITES PAYANTS

Pour l'ensemble des sites payants de la CDPFL, le CD63 nous a fourni les cumuls annuels d'entrées.

#### LE VOLCAN DE LEMPTÉGY

Pour le site du **Volcan de Lemptégy**, outre le cumul annuel d'entrées, nous avons pu obtenir des données mensuelles entre janvier 2018 et mars 2020. De plus, nous avons obtenu également, pour la réalisation d'un test sur les données Flux Vision, les données journalières sur le mois d'août 2019, avec la ventilation de ces données par origine (département, pays), pour chacune des dates du mois.

#### AUTRES SITES PAYANTS

Pour les autres sites payants, nous ne disposons que du cumul annuel des entrées fournis par le conseil départemental, pour chaque année (Figure 19 et Figure 20).

Nous avons également sollicité **Vulcania**, en novembre 2019, en même temps que Lemptégy, dans le but d'obtenir des données précises sur le mois d'août 2019 mais la Direction

avait refusé, même pour un usage de ces données uniquement pour des tests en interne, et en nous engageant à ne jamais publier aucun chiffre brut. Nous avons réitéré notre demande par la suite, en limitant notre demande à quelques dates en août 2019. Nous n'avons finalement pu obtenir, en avril 2021, un petit échantillon de 8 dates sur le mois d'août 2019, sans les origines des visiteurs.

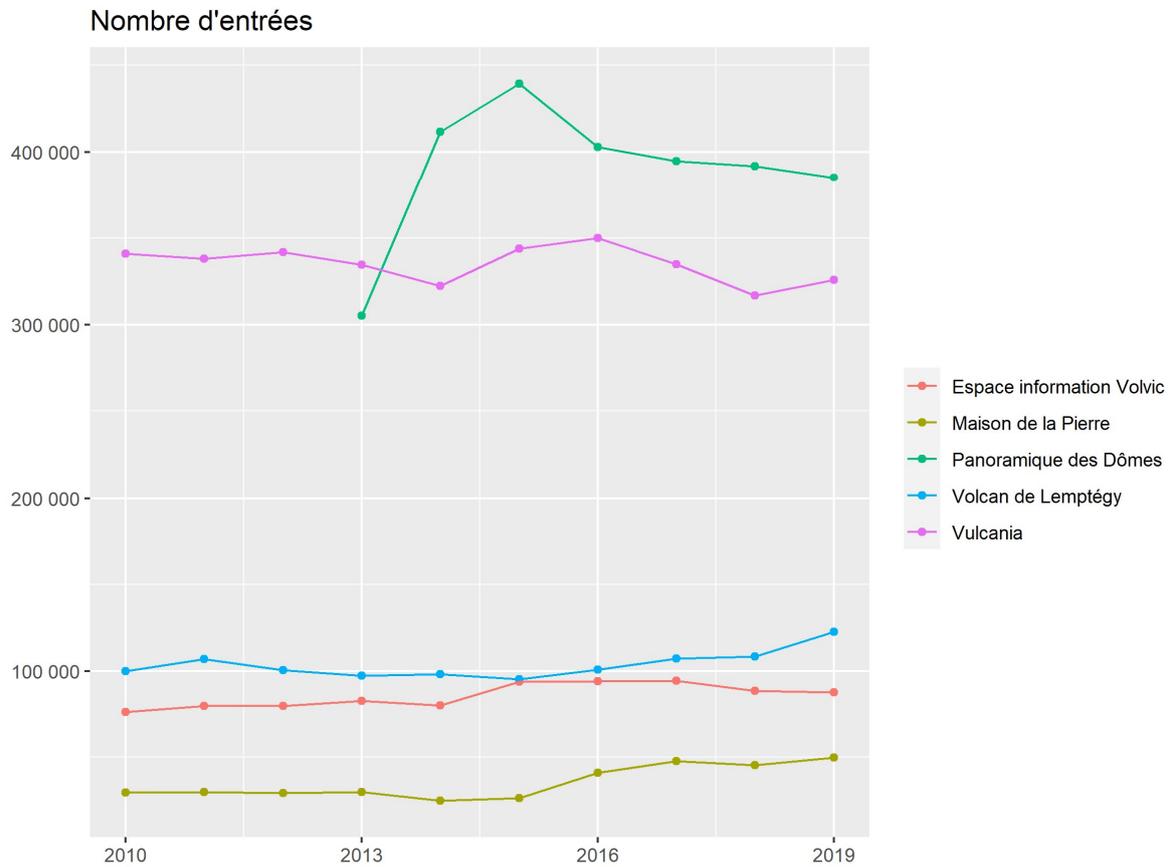


Figure 19 - Évolution du nombre annuel d'entrées sur les différents sites payants de la zone d'étude entre 2010 et 2019

## Nombre d'entrées

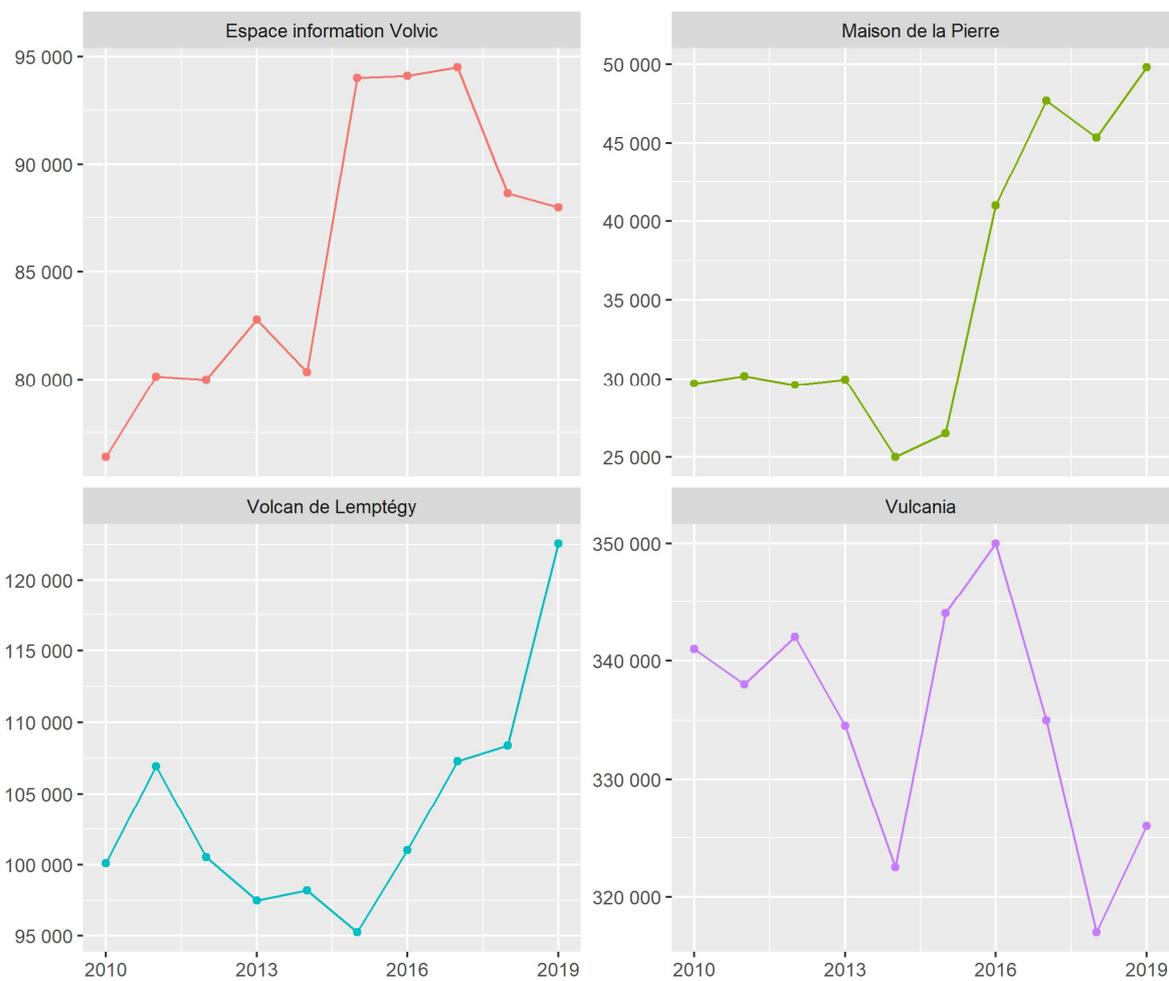


Figure 20 - Évolution du nombre annuel d'entrées par site

Enfin, sur une publication du conseil départemental (CD63, 2019) sont renseignées les calendriers d'ouverture des différents sites payants (Figure 21).

## Vulcania

Vulcania se présente comme un parc d'attractions et d'animations autour de la découverte du volcanisme mais aussi plus globalement de la planète Terre. Ce parc à thème propose de découvrir l'histoire et les paysages de la Chaîne des Puys, les mécanismes internes de notre planète et de prendre conscience de sa fragilité. Notamment au programme : expériences interactives et ludiques, sensations fortes en 3D, expositions...

### Infos pratiques :

- > période d'ouverture et horaires : 10h à 18h
- > hormis du 08/07 au 14/07 et du 24/08 au 30/08 : 10h à 18h30
- > hormis du 15/07 au 23/08 : 10h à 19h et nocturne jusqu'à 22h les 17/07, 24/07, 30/07, 31/07, 06/08, 07/08, 13/08, 14/08, 21/08, 28/08
- > en 2019, fermeture du parc du 01/01 au 19/03, les 25/03, 26/03, 01/04, 02/04 et tous les lundis et mardis entre le 02/09 et le 15/10 et du 04/11 au 31/12.

Tél. 04 73 19 70 00 – [bienvenue@vulcania.com](mailto:bienvenue@vulcania.com) – [www.vulcania.com](http://www.vulcania.com)

## Volcan de Lemptégy

Il s'agit ici de pénétrer à l'intérieur d'un vrai volcan et de découvrir cet univers surréaliste à pied ou en petit train sous la conduite d'un animateur. Ancienne carrière de pouzzolane, le volcan se présente désormais comme un cratère à ciel ouvert et devient un outil exceptionnel de vulgarisation scientifique.

### Infos pratiques : clôture de la billetterie 2h30 avant la fermeture

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Mars	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D
Avril	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	
Mai	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V
Jun	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	
Juillet	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M
Août	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S
Septembre	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	
Octobre	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J
Novembre	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	

<span style="background-color: #f08080; border: 1px solid black; padding: 2px;"> </span> 10h45 - 17h Visite guidée à pied	<span style="background-color: #4682b4; border: 1px solid black; padding: 2px;"> </span> 9h30 - 18h Visite guidée en petit train	<span style="background-color: #ffff00; border: 1px solid black; padding: 2px;"> </span> 9h30 - 18h Visite guidée en petit train ou à pied (au choix)
<span style="background-color: #f08080; border: 1px solid black; padding: 2px;"> </span> 10h45 - 17h Visite audioguidée à pied	<span style="background-color: #4682b4; border: 1px solid black; padding: 2px;"> </span> 9h30 - 18h Visite audioguidée à pied	<span style="background-color: #008000; border: 1px solid black; padding: 2px;"> </span> 9h - 19h Visite guidée en petit train ou à pied (au choix)
<span style="background-color: #f08080; border: 1px solid black; padding: 2px;"> </span> 10h - 17h30 Visite audioguidée à pied	<span style="background-color: #f08080; border: 1px solid black; padding: 2px;"> </span> 10h30 - 17h30 Visite guidée en petit train (à pied en cas de neige ou gel)	
<span style="background-color: #90ee90; border: 1px solid black; padding: 2px;"> </span> 9h30 - 18h30 Visite guidée en petit train ou à pied (au choix)		

Tél. 04 73 62 23 25 – [info@auvergne-volcan.com](mailto:info@auvergne-volcan.com) – [www.auvergne-volcan.com](http://www.auvergne-volcan.com)

## Grotte de la Pierre de Volvic

La grotte de la Pierre de Volvic permet un voyage au cœur d'une coulée de lave et une immersion au temps des tailleurs de pierre dans les années 1900. Cinq espaces muséographiques permettent d'appréhender l'activité emblématique de notre région qu'est la taille de la pierre.

### Infos pratiques : dernière visite 1h15 avant la fermeture

- > du 12 février au 6 avril : ouvert mardi, mercredi, jeudi de 11h à 13h et de 14h à 17h.  
Ouverture exceptionnelle Trail de Vulcain samedi 2 mars de 14h à 17h
- > du 7 avril au 7 juillet : ouvert tous les jours 11h à 18h
- > du 8 juillet au 24 août : ouvert tous les jours de 10h à 19h
- > du 25 août au 22 septembre : fermé le samedi et ouvert de 11h à 13h et 14h à 18h
- > du 23 septembre au 1<sup>er</sup> novembre : fermé le samedi et ouvert de 14h à 18h

Tél. 04 73 33 56 92 – [www.maisondelapierre-volvic.fr](http://www.maisondelapierre-volvic.fr)

Figure 21 - Calendrier d'ouverture des sites payants de CPFL - source : (CD63, 2019)

---

#### 2.1.4. AYDAT

N'ayant pu trouver aucune information quant à la fréquentation d'Aydat, dans aucune publication, nous avons pris contact avec le service tourisme de Mond'Arverne, pour essayer de savoir s'il existait des données sur la fréquentation d'Aydat. Nous avons pu échanger en visio avec Carole Carlet-Bonnet, le 18 janvier 2021. Nous avons appris que les seules données dont ils disposaient étaient les données d'entrées de l'office de tourisme d'Aydat. Il nous a également été signalé qu'il existait des comptages ponctuels du nombre de baigneurs réalisés par les sauveteurs pendant l'été, ainsi que données concernant les activités de nautisme gérées par le club nautique.

Nous avons pu obtenir les données de fréquentation de l'office de tourisme sur l'année 2019. En revanche, nous n'avons pas eu accès aux comptages réalisés par les sauveteurs. Toutefois, durant notre échange, il a été question d'un nombre maximal de personnes présentes, au même moment, sur le site d'Aydat durant l'été de l'ordre de 2000 personnes.

---

#### 2.1.5. AUTRES DONNEES

---

##### ENQUETE VISITEURS

Les données de l'enquête visiteurs réalisées dans le cadre du projet Valo-Sites, entre mai 2019 et février 2020, ont été utilisées pour différents calculs, au cours de la tâche T1B, en premier lieu pour estimer la récurrence des visites d'un touriste sur la CPFL et la fréquence annuelle des visites d'un excursionniste. D'autre part, les données Flux Vision qualifiant les touristes ont été confrontées aux données de cette enquête.

Plus de mille visiteurs ont été interrogés (1011) sur toute la période et sur les différents sites de CPFL, représentant près de 3000 visiteurs (2894). Un échantillonnage par quotas sur les sites et les périodes de l'année, c'est-à-dire au prorata de la fréquentation connue sur les sites dénombrables (sites payants, éco-compteurs), et en répartissant les dates d'enquête sur les mois de l'année en fonction de l'affluence relative sur chaque mois, a été choisi pour obtenir un échantillon représentatif de la population des visiteurs (touristes et excursionnistes).

Les données ont finalement été redressées en fonction des probabilités d'inclusion dans l'échantillon et redressées par calage sur les périodes de l'année, les quotas n'ayant pu être respectés parfaitement, notamment en raison du Covid-19 (Marsat, 2022).

Enfin, une expérimentation permettant de tester des dispositifs allégés d'enquête a été menée dans le cadre du stage d'Orane Boucher en 2021, et a notamment permis d'acquérir des données via une enquête auto-administrée en ligne (via un affichage sur les sites avec QR code), en particulier sur le site de Lemptégy, où le dispositif a très bien fonctionné (Boucher, 2021).

---

##### AUTRES DONNEES UTILISEES

Enfin d'autres données ont été collectées au cours du projet :

- Des données de contexte permettant de réaliser les imputations et extrapolations sur les séries incomplètes : données météorologiques (Données issues de la plateforme INRAE CLIMATIK <https://agroclim.inrae.fr/climatik/>, gérée par l'unité AgroClim d'Avignon), des données calendaires, des dates d'évènements récupérées sur des sites dédiés : sportifs (ex. Figure 22), culturels, sorties... ;
- Des données de fréquentations touristiques fournies par d'autres sources : enquête INSEE / Baromètre CRT AURA (Figure 23), METT / CRT AURA (Figure 24), enquête SOFRES Puy de Dôme, chiffres-clé CD63, données OT (Mond'Arverne, Clermont-Auvergne-Métropole, Terra Volcana) ;
- Données sur l'hébergement touristique (INSEE), la population des communes, et les données de population carroyées (INSEE, 2019) ;
- Données Flux Vision du Cirque de Navacelles Grand Site Occitanie, transmises par le CRT Occitanie, pour éventuellement réaliser les mêmes tests sur un autre site naturel, en les confrontant aux données du Grand Site de France (éco-compteurs, comptages routiers/parking, navettes ; données non-récupérées).

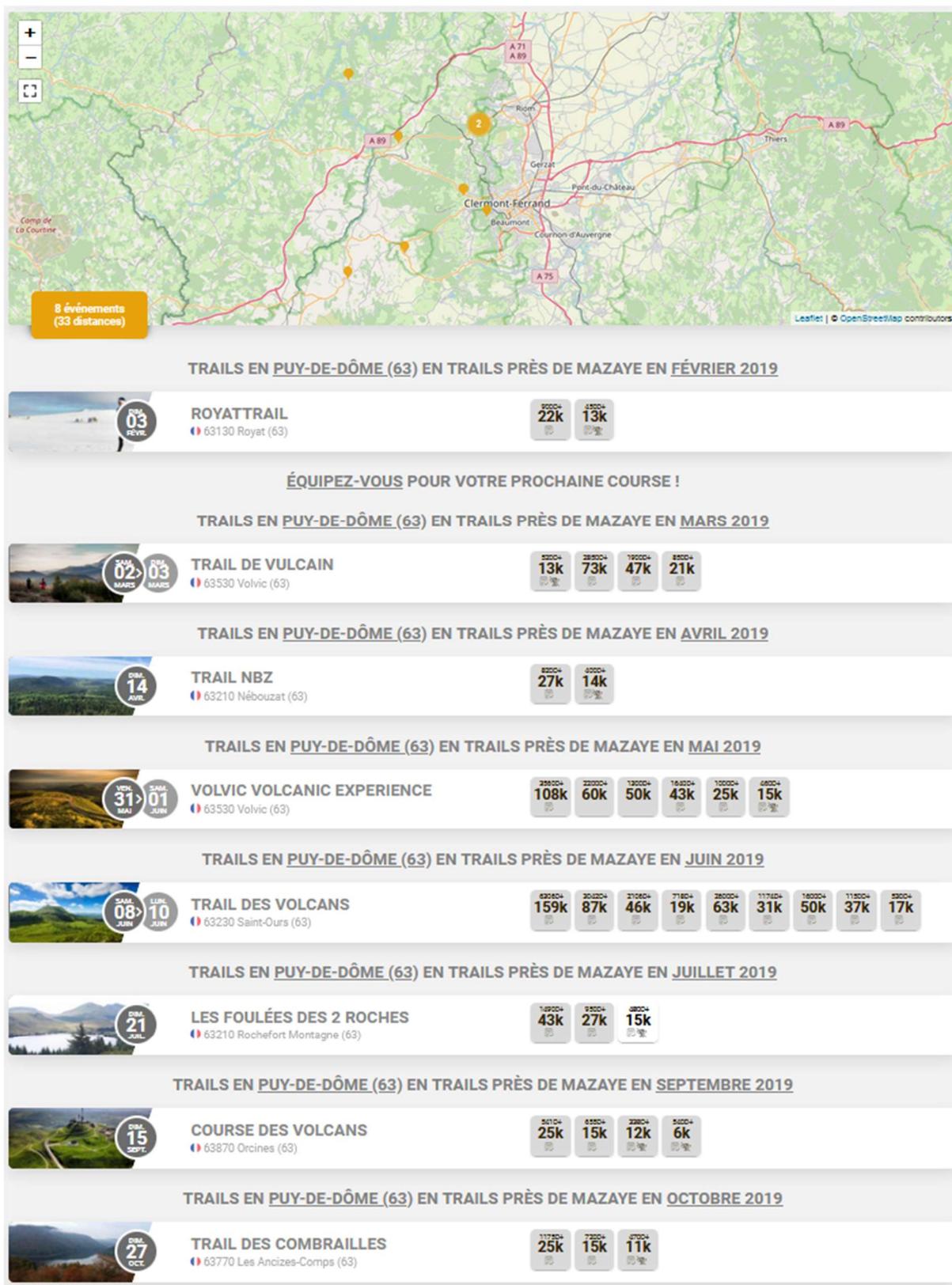


Figure 22 - Trails dans ou proches de la zone CPFL en 2019 (Capture d'écran du site [www.betrail.run](http://www.betrail.run))

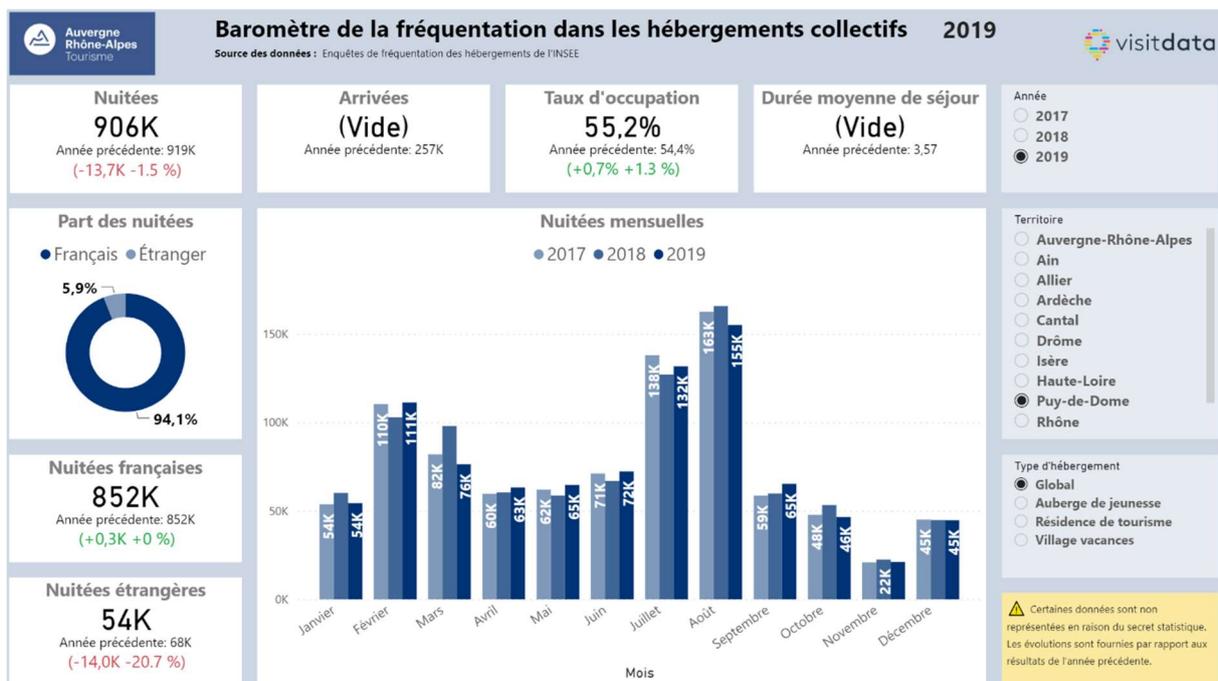


Figure 23 - Baromètre de la fréquentation dans les hébergements collectifs en 2019 dans le Puy de Dôme – CRT Auvergne-Rhône-Alpes (source : enquête Insee sur la fréquentation des hébergements collectifs d’Auvergne-Rhône-Alpes)

	Nombre de lits	Nombre de nuitées <i>Estimation</i>	Retombées économiques <i>Estimation</i>
<b>UNESCO Chaîne des Puy</b>			
Hôtels	1 133 <span style="color: red;">53%</span>	207 200 <span style="color: blue;">97%</span>	30 400 000 € <span style="color: blue;">97%</span>
Campings et PRL	2 293 <span style="color: orange;">64%</span>	75 600 <span style="color: orange;">64%</span>	4 340 000 € <span style="color: orange;">64%</span>
Résidences de tourisme et résidences hôtelières			
Villages vacances, villages de gîtes et maisons familiales			
Centres de vacances, auberges de jeunesse...	274 <span style="color: blue;">100%</span>	28 300 <span style="color: blue;">100%</span>	1 869 000 € <span style="color: blue;">100%</span>
Gîtes d'étape, gîtes de groupes et refuges	263 <span style="color: orange;">86%</span>	18 400 <span style="color: orange;">86%</span>	1 542 000 € <span style="color: orange;">86%</span>
Chambres d'hôtes et hébergements insolites	223 <span style="color: blue;">97%</span>	20 900 <span style="color: blue;">97%</span>	1 346 000 € <span style="color: blue;">97%</span>
Meublés, gîtes et HLL	732 <span style="color: orange;">80%</span>	78 800 <span style="color: blue;">100%</span>	5 288 000 € <span style="color: blue;">100%</span>
<b>HÉBERGEMENT MARCHAND</b>	<b>4 918</b>	<b>429 200</b>	<b>44 785 000 €</b>
<b>RÉSIDENCES SECONDAIRES</b>	<b>10 412</b>	<b>199 400</b>	<b>5 541 000 €</b>

Figure 24 - Module d'Évaluation du poids Touristique des Territoires (METT) - Zone UNESCO en mars 2021 (source : Base APIDAE, INSEE, Euroeka, Kantar TNS, Acoos / Urssaf)

## 2.2. TRAITEMENT DES DONNEES D'ECO-COMPTEURS

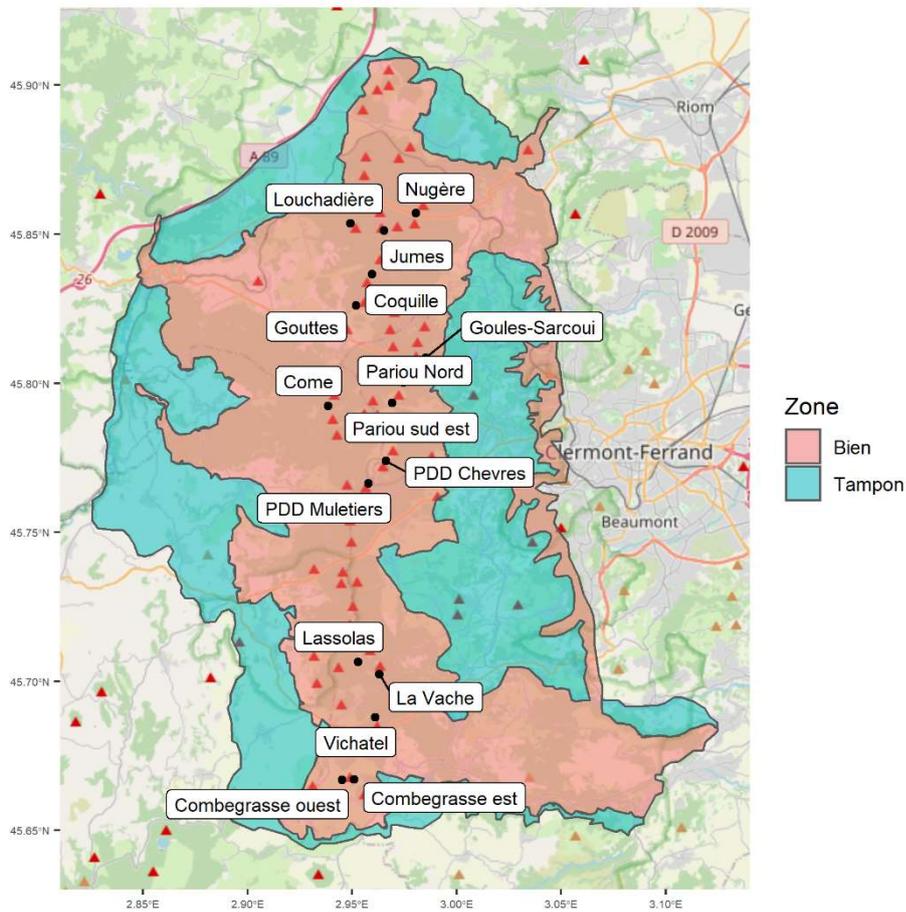


Figure 25 - Carte des éco-compteurs utilisés

### 2.2.1. DIAGNOSTIC DES SERIES

Les données d'éco-compteurs ont été exploitées principalement au grain journalier. Les données infra-horaires ont été utilisées dans l'analyse exploratoire pour mieux comprendre les flux entre puy, notamment sur la zone Puy de Dôme – Pariou.

Une première analyse des séries a consisté à identifier des anomalies sur les séries IN et OUT des 16 éco-compteurs, étape indispensable compte tenu des problèmes inhérents à l'acquisition de ces données (dysfonctionnements, arrêt, vandalisme...). Plusieurs étapes de diagnostic ont permis de repérer des données ou séquences potentiellement anormales, nécessitant une analyse complémentaire, puis éventuellement des corrections.

La détection des anomalies potentielles consiste à repérer des motifs particuliers dans les séquences, tels que :

- Des pics anormalement élevés, pouvant correspondre à un évènement (courses, groupes...) ou d'autres causes (troupeau ?),
- Des inversions apparentes de sens IN et OUT,

- Des incohérences par site, entre les flux entrant et sortant cumulés sur une journée,
- Enfin, et surtout, des séquences anormalement basses, pouvant correspondre à des dysfonctionnements du capteur.

L'étape suivante revient à automatiser l'étape d'identification, en développant un algorithme, qui soit à la fois sensible, pour détecter tous les motifs atypiques, et spécifique, pour ne pas inclure des données ou séquences normales. Cet algorithme permettra le filtrage de ces motifs, par exemple pour les corriger ou les ré-estimer.

---

## ÉTAPE 1 : IDENTIFICATION DE PICS

Plusieurs algorithmes ont été testés afin d'identifier des valeurs anormalement élevées, algorithmes basés sur des quantiles ou des moyennes mobiles.

L'algorithme retenu consiste en un triple-filtre discriminant des valeurs vérifiant les 3 conditions suivantes :

- Valeurs supérieures à un seuil de 150 passages journaliers ;
- Valeurs supérieures à 3 fois la valeur d'une moyenne mobile centrée sur 15 jours (élaguée à 10%) ;
- Valeurs supérieures à 2 fois le quantile 98.5% de la série.

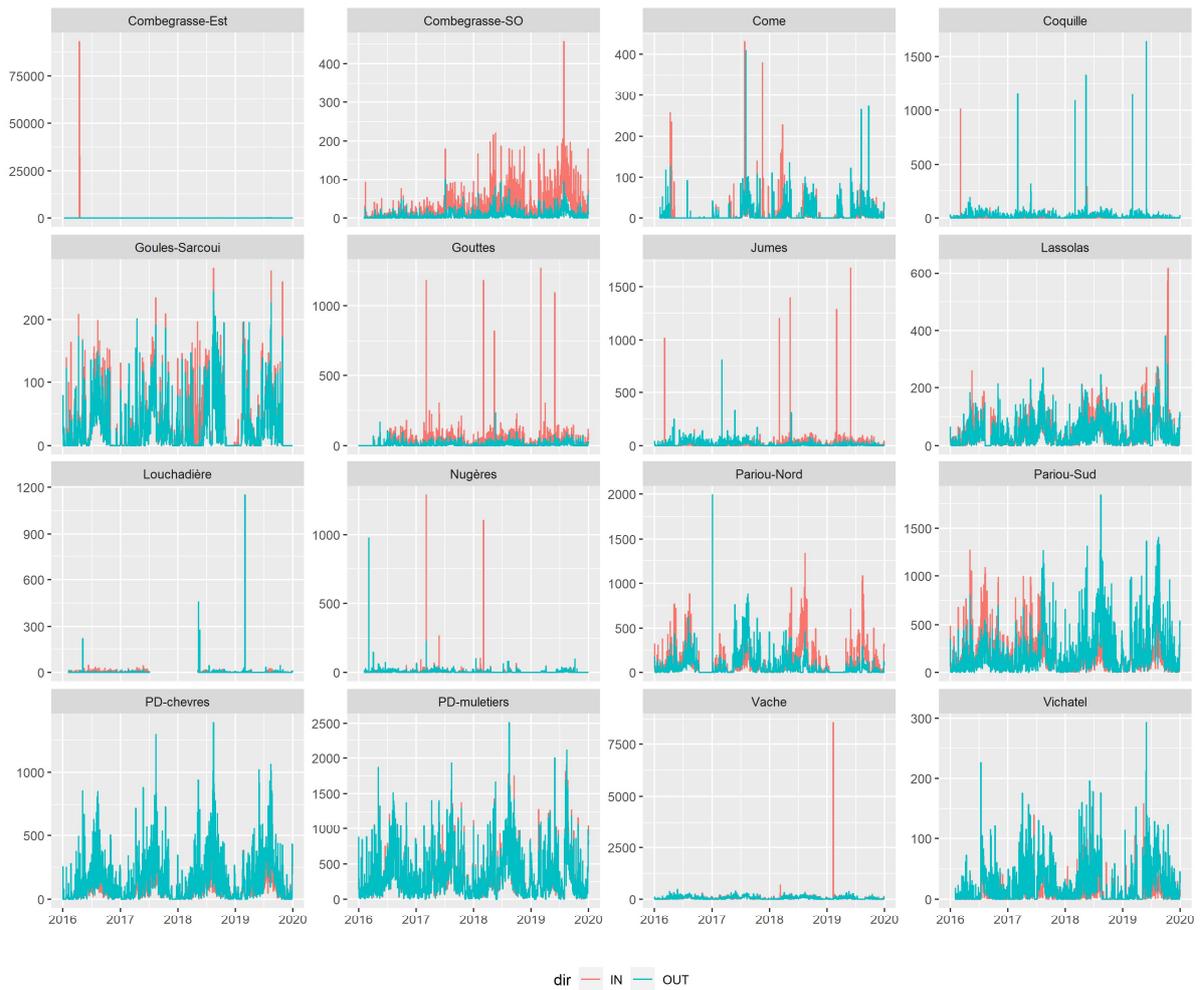


Figure 26 - Séries brutes des nombres de passages journaliers dans les deux directions (IN et OUT) pour les années 2016 à 2019

L'application du filtre sur les données brute (Figure 26) permet d'identifier ces valeurs anormalement élevées, représentées par des points sur la Figure 27, et de les extraire de la série, afin d'obtenir une série sans accident (Figure 28), donc plus facilement modélisable.



**Figure 27 - Détection des pics anormaux (matérialisés par un point) avec un triple filtre sur les séries brutes des nombres de passages journaliers dans les deux directions (IN et OUT) pour les années 2016 à 2019**

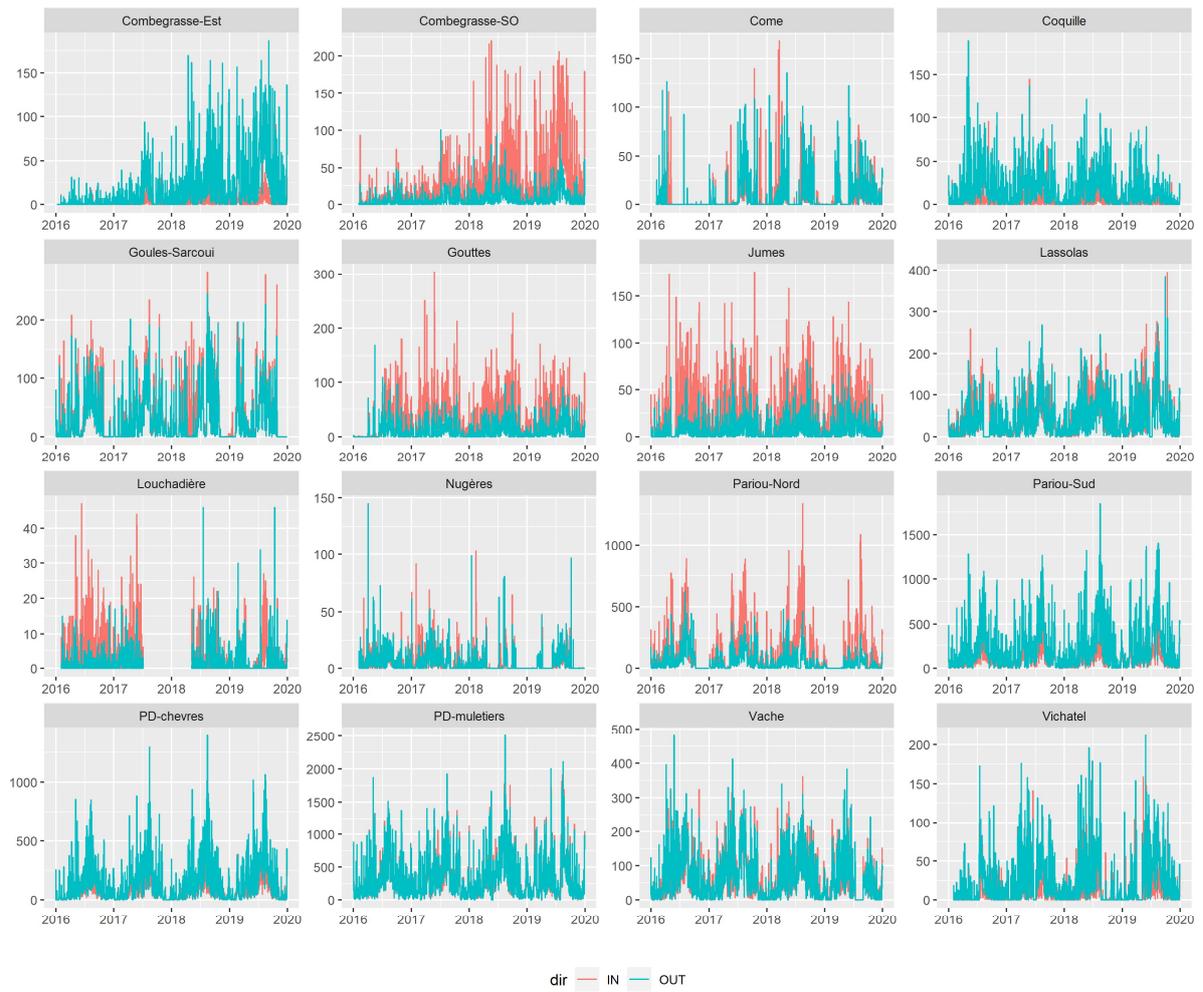


Figure 28 - Séries des nombres de passages journaliers dans les deux directions pour les années 2016 à 2019 sans les pics anormaux

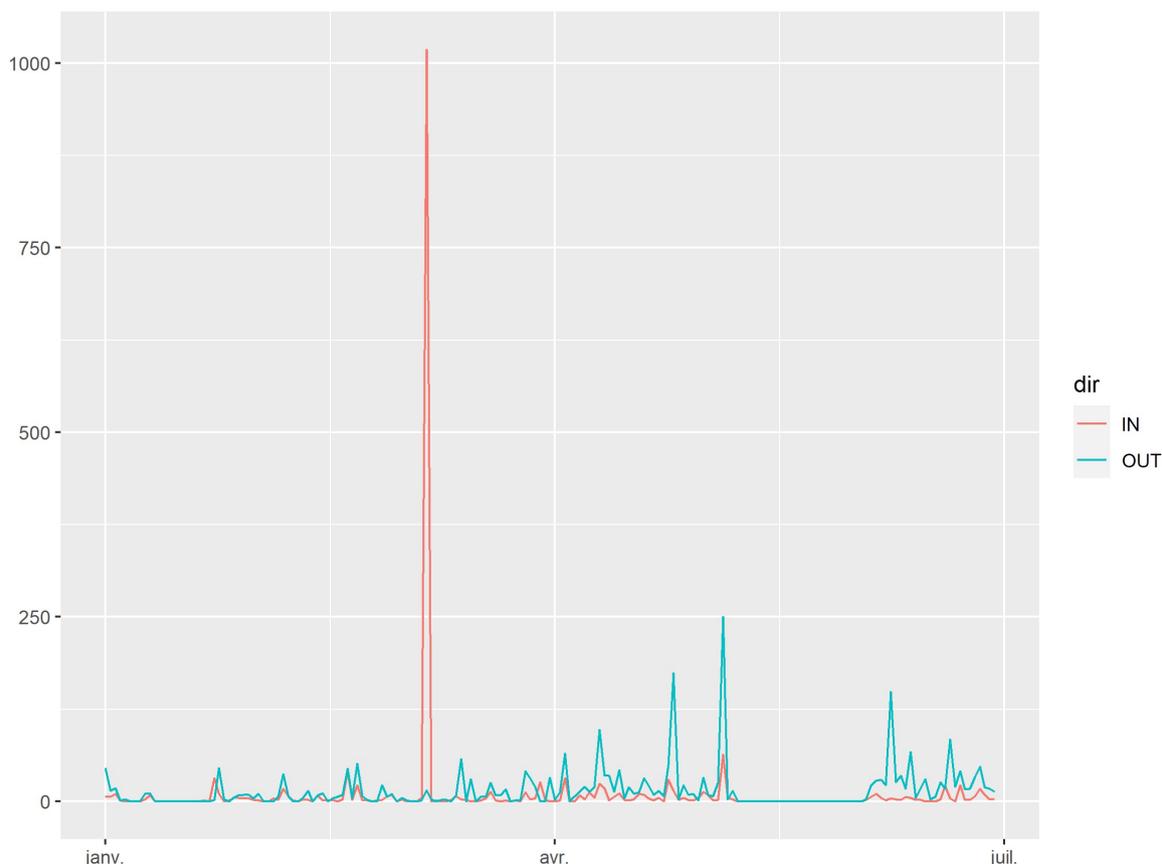


Figure 29 - Éco-compteur de Jumes (début 2016)

Pour illustrer cela, sur la Figure 29 représentant les séries de l'éco-compteur de Jumes, au début de l'année 2016, on peut voir un pic très élevé correspondant à un événement sportif : le trail du Vulcain, qui a eu lieu le 5 mars 2016 (remarque : on peut également voir une période anormalement basse sur la période fin mai - début juin). Ces pics devraient donc être confrontés à un calendrier des événements susceptibles de générer une forte fréquentation, afin de distinguer les vrais pics de fréquentation des pics anormaux (troupeaux ?).

## ÉTAPE 2 : SEQUENCES ANORMALEMENT BASSES

Plus gênant, on constate sur la plupart des éco-compteurs, des séquences anormalement basses, correspondant probablement à un dysfonctionnement du capteur. En observant, par exemple, les courbes lissées superposées des différentes années (Figure 30), on peut facilement voir, pour différents éco-compteurs, des séquences anormalement basses, parfois quasi-nulle sur plusieurs jours (par exemple au printemps 2019 sur Louchadière, aux printemps 2018 et 2019 sur Nugères, et en juillet 2019 sur le sentier des muletiers et sur Vache).

Pour détecter automatiquement ces séquences anormalement basses, plusieurs indicateurs ont été testés. L'idée générale est de comparer la valeur du jour à une valeur attendue à la même date de l'année, et de trouver un seuil adéquat permettant de discriminer les séquences réellement trop basses, et non de creux conjoncturels, c'est-à-dire ponctuels et sans doute explicables.

Pour détecter ces séquences anormalement basses, l'idée est de comparer la donnée journalière à un quantile des données de la série à la même date, ou à la même période, sur l'ensemble des années disponibles. Une valeur qui serait inférieure à une très petite fraction de ce quantile serait donc jugée anormale. Plusieurs quantiles (médiane, quantile à 75%), et plusieurs fractions (seuils) ont été testés. Finalement, les données ont été filtrées à partir d'une courbe médiane, en fixant le seuil à  $1/10^{\text{ème}}$  de cette courbe. Les données inférieures à cette courbe limite sont considérées comme anormalement basses. Les séquences filtrées ont ensuite été étendue à chaque extrémité (via une fenêtre mobile) afin d'avoir des séquences complètes sans discontinuité, considérant que lorsque l'éco-compteur dysfonctionne, toutes les données sont erronées jusqu'à la remise en état du capteur, et que la dégradation des données est peut-être progressive au début.

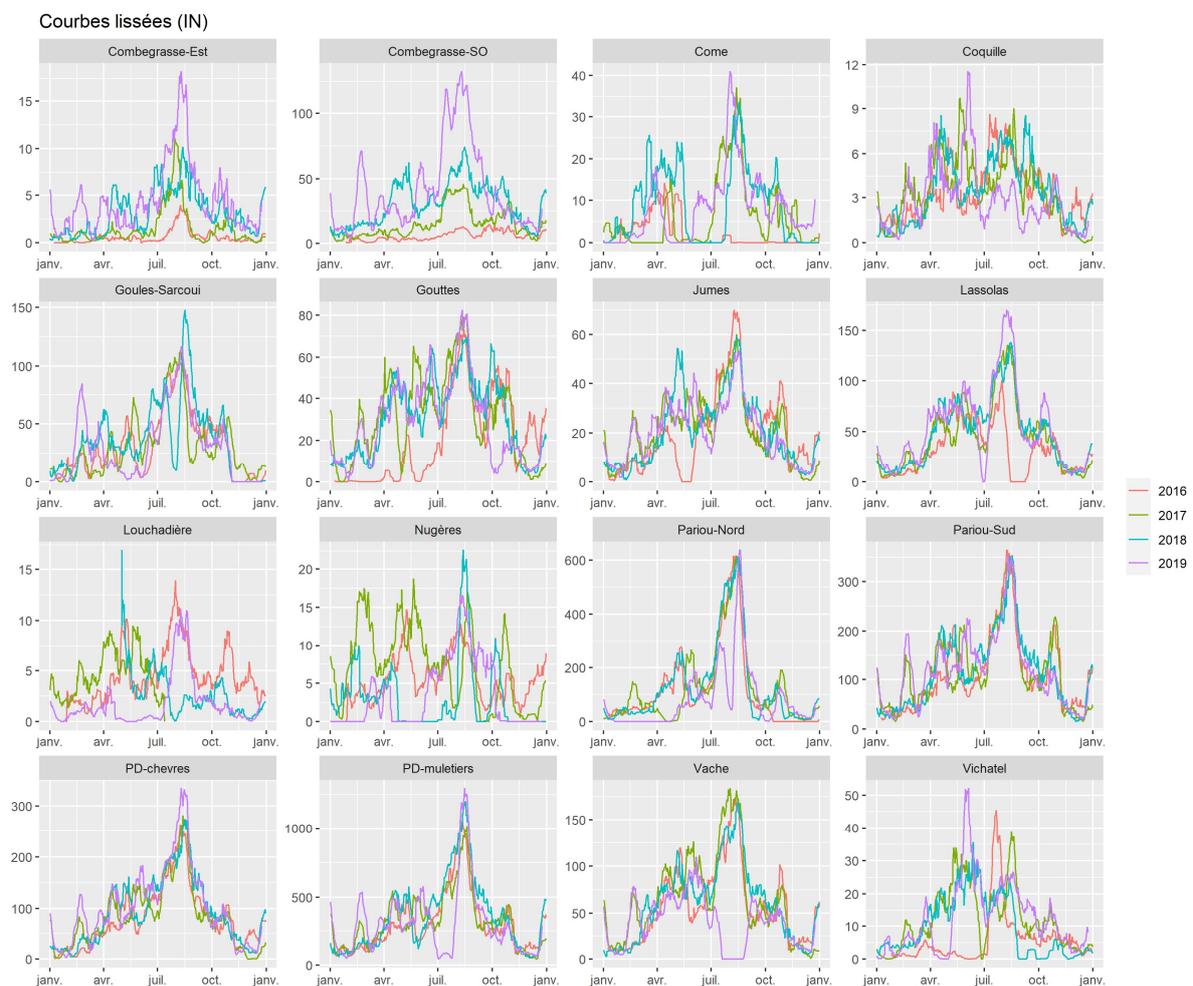


Figure 30 - Séries lissées (par année)

On peut voir, sur Figure 31, les séquences filtrées (sur fond bleu) et, sur Figure 32, l'importance de ces séquences (bleu clair) sur l'ensemble des séries. Il faut noter que ce filtrage est excessif en plusieurs endroits, notamment sans doute en début de série sur les éco-compteurs de Combegrasse, en raison de la nouveauté du circuit, ou encore sur de très courtes séquences en divers endroits, pouvant peut-être s'expliquer par une mauvaise météo sur plusieurs jours.

Passages IN + filtre Médiane

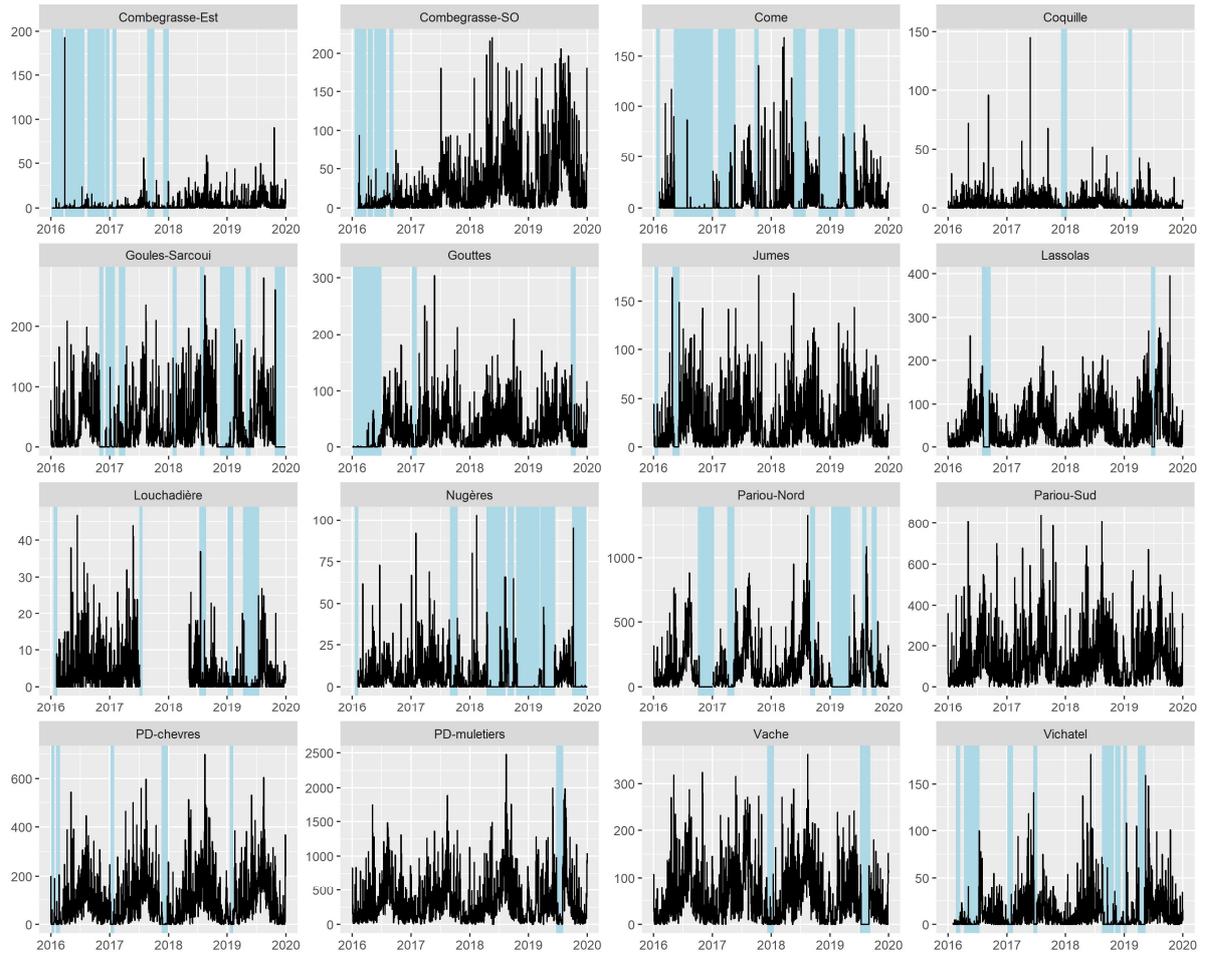
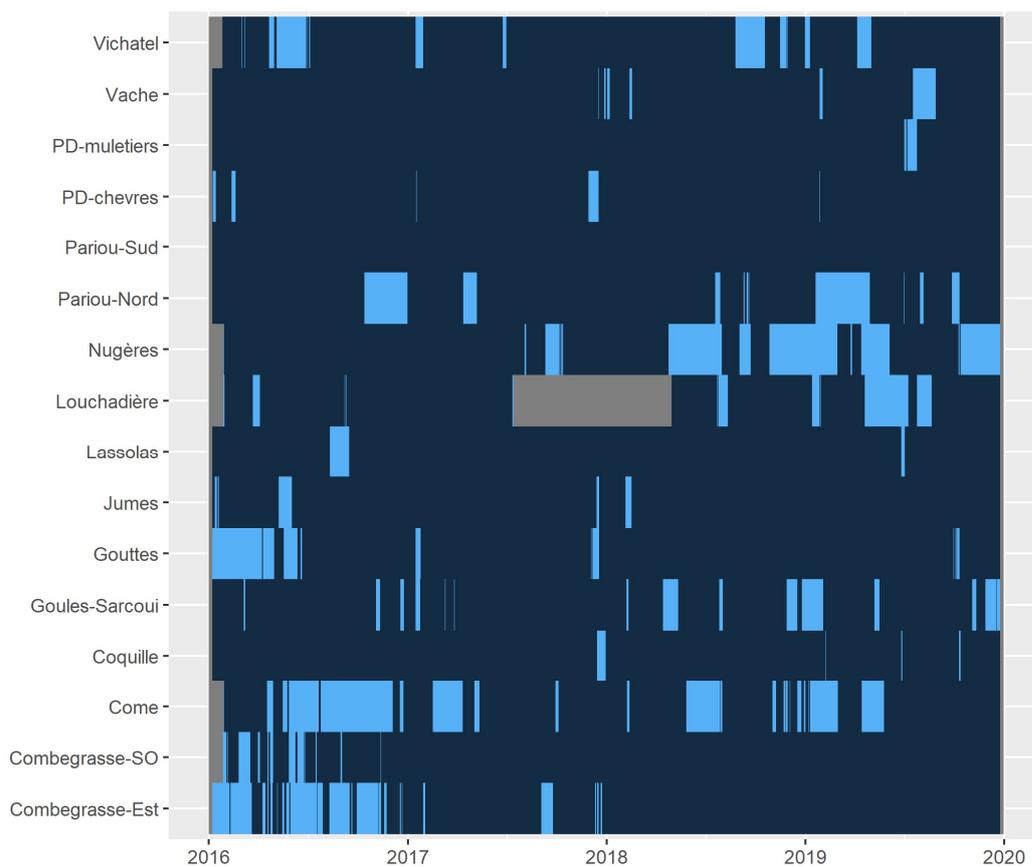


Figure 31 - Détection des séquences anormalement basses (filtre médiane)



**Figure 32 - Répartition des séquences anormalement basses (bleu clair)**

On peut voir sur Figure 33, sur l'éco-compteurs de Pariou-Nord, des séquences clairement anormales, à la fin 2016 et au début de 2019, ainsi que d'autres séquences basses peut-être plus explicables.

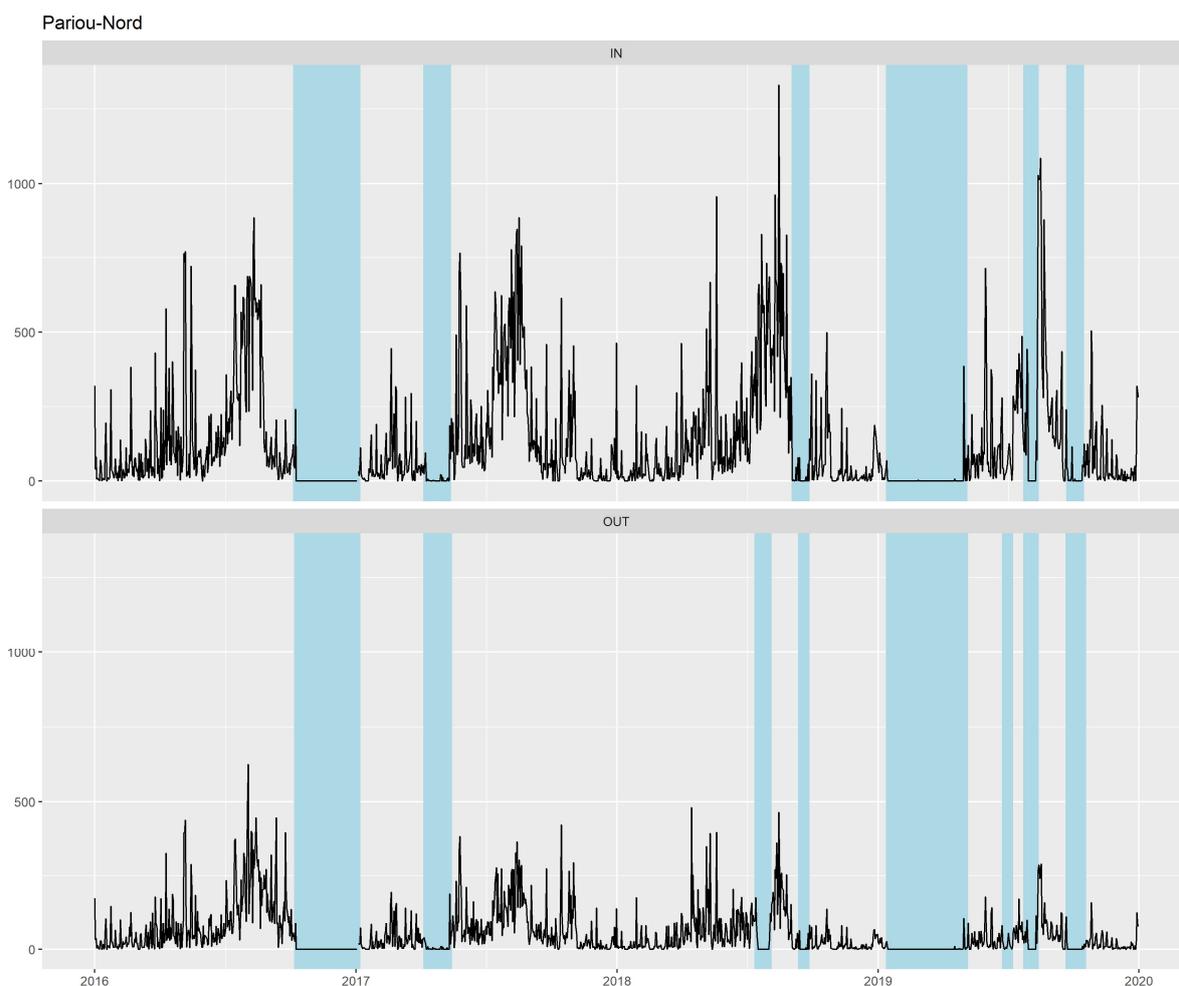


Figure 33 - Séries IN et OUT de l'éco-compteur Pariou-Nord avec séquences anormalement basses (fond bleu clair)

### ÉTAPE 3 : INVERSION DES FLUX IN ET OUT

De la même façon que précédemment, en regardant les séries des données mensuelles (Figure 34), on peut facilement voir que sur certains éco-compteurs, les courbes IN et OUT semblent s'inverser au cours du temps, ce qui semble difficilement explicable a priori. On peut notamment le voir sur la courbe de l'éco-compteur de Pariou-Nord (Figure 36), avec des écarts importants entre le IN et le OUT, qui s'inversent en 2017 par rapport aux autres années.

Une autre façon de détecter ces inversions consiste à calculer un indicateur d'inversion mesurant la moyenne du signe des écarts  $IN - OUT$ . Pour les éco-compteurs où ces inversions sont détectables, c'est-à-dire lorsque les écarts entre IN et OUT sont importants, l'indicateur sera proche de 1 ou de -1. En cas d'inversion, il changera brutalement de signe, comme on peut par exemple le voir pour l'éco-compteur de Jumes entre fin 2017 et début 2018, ou encore pour celui de Pariou-Sud au milieu de l'année 2017 (Figure 35).

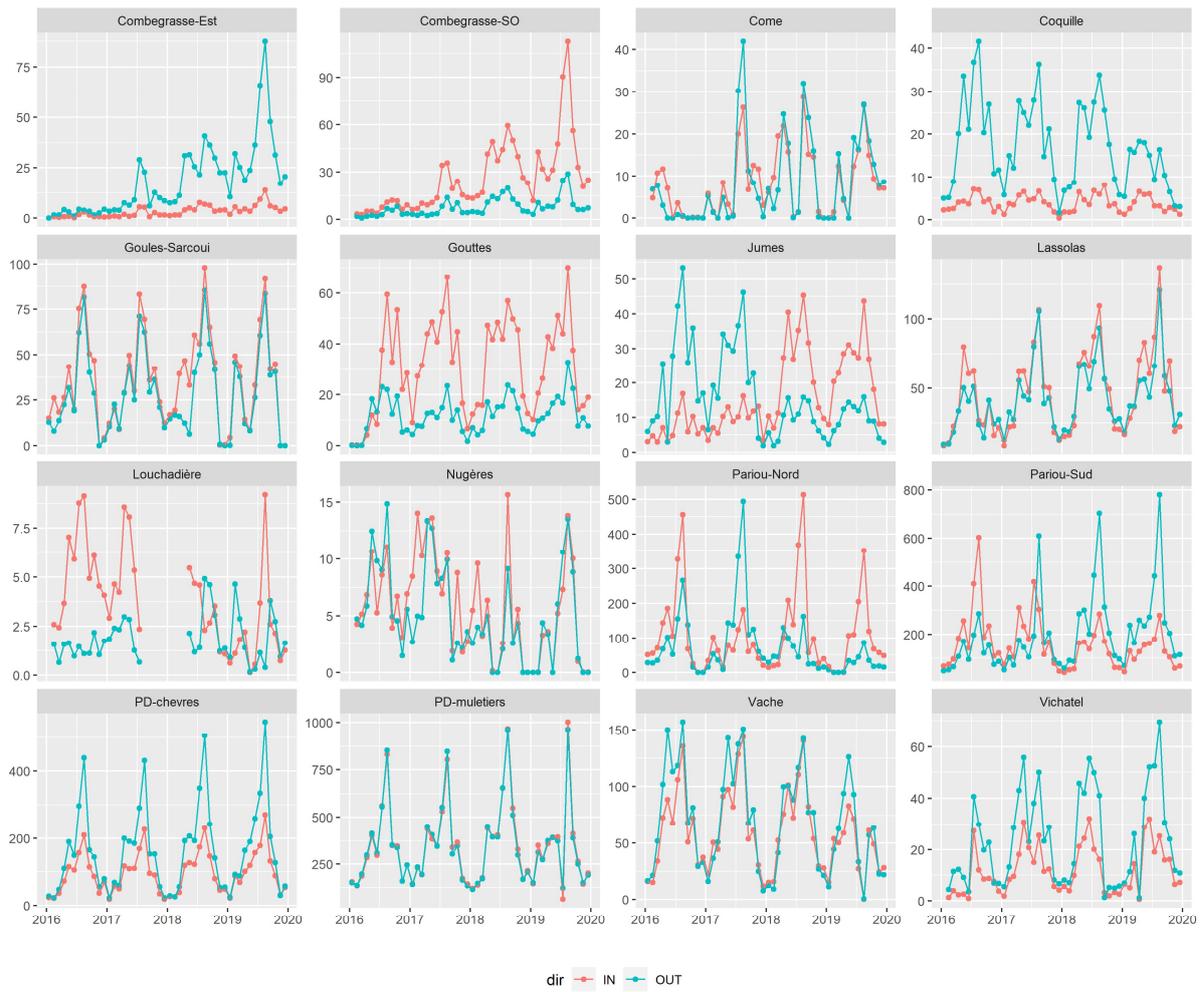


Figure 34 - Séries des moyennes mensuelles

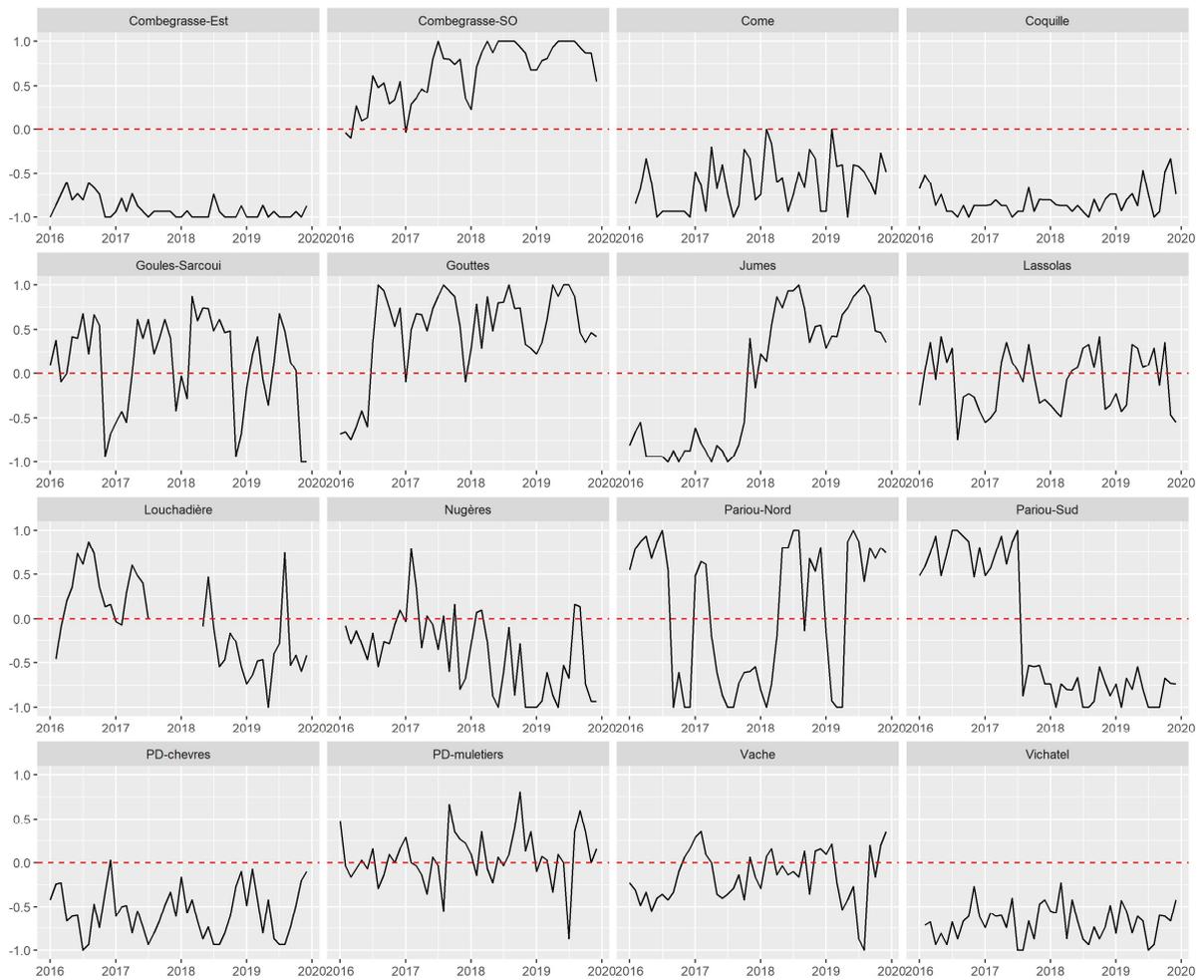


Figure 35 - Indicateur d'inversion

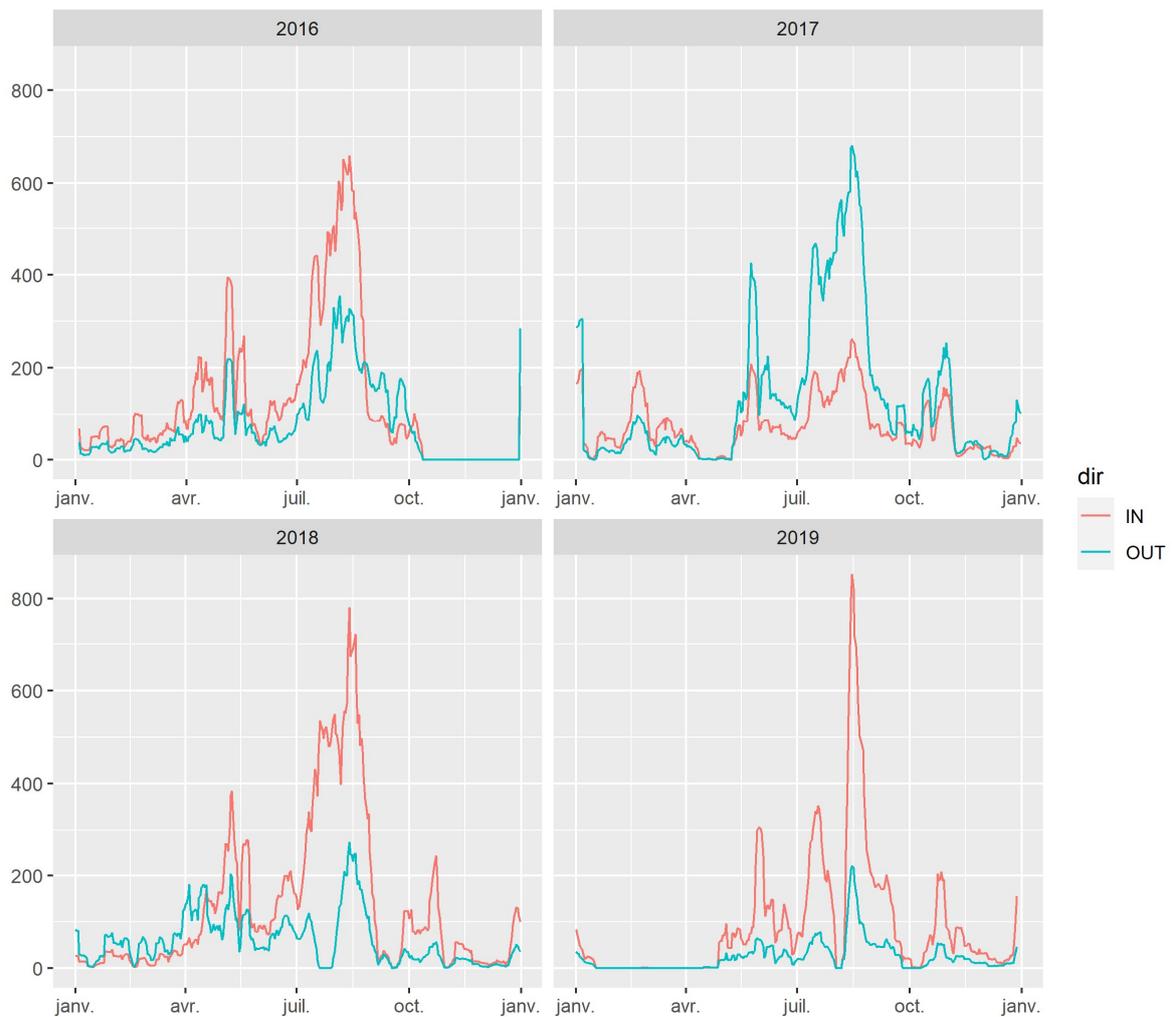


Figure 36 - Séries IN et OUT de l'éco-compteur Pariou-Nord (par année)

#### ÉTAPE 4 : INCOHERENCE DES CUMULS IN ET OUT

En l'absence de pertes, c'est-à-dire si tous les visiteurs d'un puy passent effectivement par les éco-compteurs, on doit s'attendre à avoir, sur une journée, égalité entre l'ensemble des flux entrant et sortant, soit  $\sum_{EC} IN = \sum_{EC} OUT$ .

Par exemple, si l'on considère les deux éco-compteurs du Pariou, on devrait avoir sur une journée :

$$\begin{aligned} \text{Pariou} - \text{Nord}_{IN} + \text{Pariou} - \text{Sud}_{IN} \\ = \text{Pariou} - \text{Nord}_{OUT} + \text{Pariou} - \text{Sud}_{OUT} \end{aligned}$$

De même, pour le Puy de Dôme, après ajout des données aller/retour du Panoramique des Dômes :

$$\begin{aligned} PD - \text{muletiers}_{IN} + PD - \text{chèvres}_{IN} + PDD_{\text{aller}} = \\ PD - \text{muletiers}_{OUT} + PD - \text{chèvres}_{OUT} + PDD_{\text{retour}} \end{aligned}$$

On peut donc calculer un nouvel indicateur mesurant l'écart journalier entre les flux entrant et sortant. Cet indicateur (Figure 37) se traduit le plus souvent par des écarts limités, proportionnés à la fréquentation du site et de la période (ex. Combegrasse), mais qui peuvent être parfois très importants, dépassant par exemple 500, voire 750, en valeur absolue, sur le Puy de Pariou.

On peut noter les écarts importants et systématiquement positifs sur le Puy de Dôme, pouvant peut-être s'expliquer partiellement par les parapentistes, qui empruntent le Panoramique des Dômes, en montant, et qui redescendent par les airs. D'autres écarts importants mais de signe variable, comme sur Vache-Lassolas, peuvent peut-être s'expliquer par les circuits alternatifs rejoignant le puy de Lassolas, sans passer par les éco-compteurs.

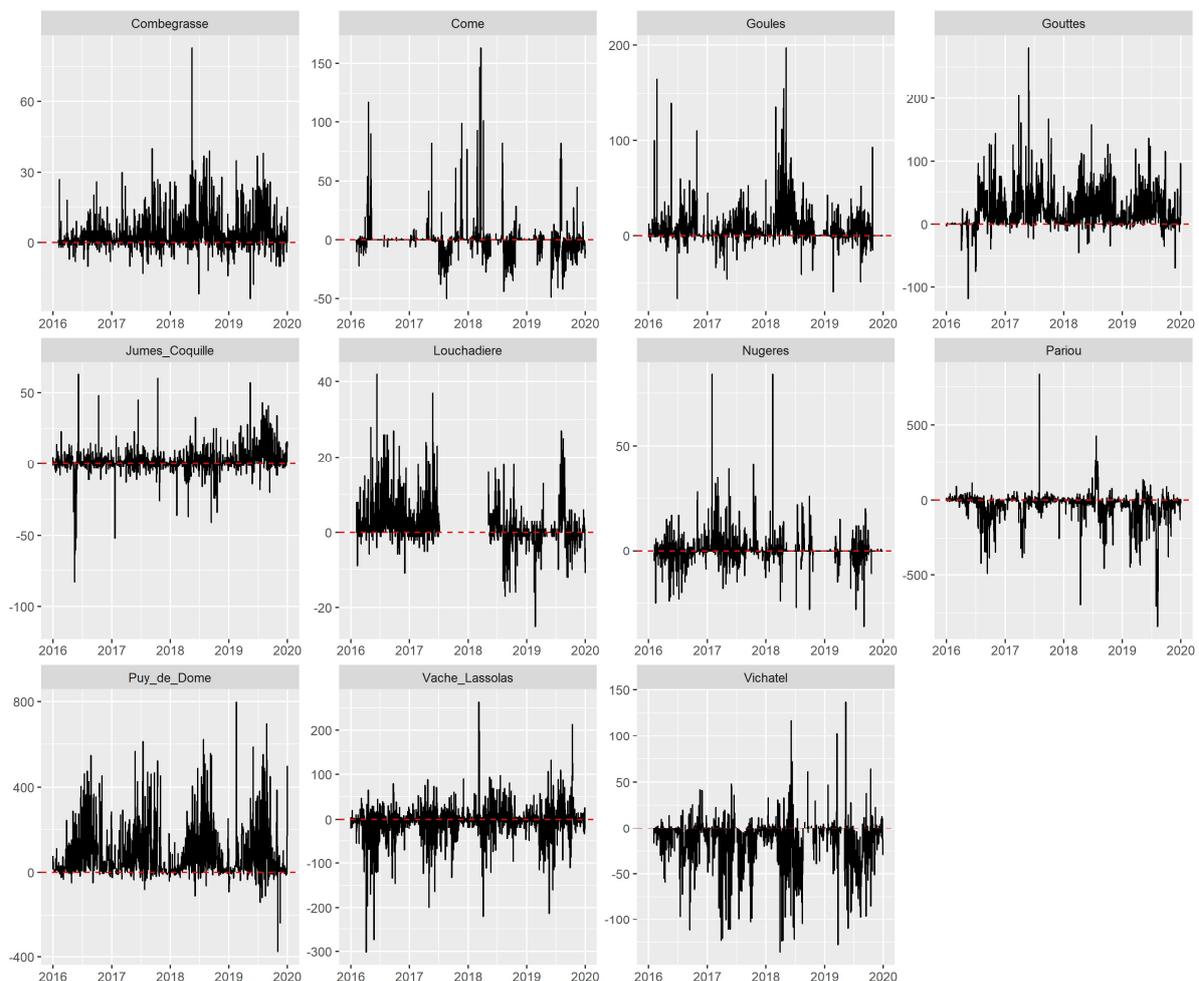


Figure 37 - Série des écarts de cumul IN - OUT par site

## 2.2.2. TRAITEMENTS

Ayant identifié dans les séries de données les principaux motifs pouvant correspondre à diverses anomalies : données extrêmes ou aberrantes, séquence de données anormalement basses, incohérences IN/OUT dans le temps (inversion des rapports) et sur une journée (cumuls

inégaux), il s'agit désormais de voir dans quelle mesure l'on pourrait corriger certaines de ces données douteuses.

En utilisant des algorithmes ou des indicateurs développés pour détecter automatiquement ces motifs, et éventuellement les filtrer, des corrections (a minima) ont été réalisées pour les anomalies les plus flagrantes.

Les séquences anormalement basses ont été filtrées, puis imputées. Pour ce faire, plusieurs algorithmes et plusieurs modèles ont été testés. Enfin, après imputation, les pics, préalablement filtrés, ont été réinjectés dans les séries.

---

## NETTOYAGE DES MOTIFS ABERRANTS

Une partie de ces données, identifiées à partir des indicateurs proposés plus haut, a donc été corrigée, ou bien masquée, pour être ensuite ré-estimée, c'est-à-dire imputée, comme s'il s'agissait de données manquantes. Il a été fait le choix de ne pas corriger certains motifs, tels que les différences de cumuls, en raison d'un manque de connaissances a priori sur les processus ayant pu les générer (ex. passages multiples devant la cellule, passages hors-sentier, troupeaux...). Toutefois, il serait essentiel d'enrichir la base avec un recensement, et si possible un historique, de tous les phénomènes pouvant impacter la mesure faite par les éco-compteurs (événements sportifs ou culturels, groupes scolaires...). Il serait également important de réaliser un test de la fiabilité des mesures, notamment quant aux passages simultanés.

---

## IMPUTATIONS

Plusieurs approches d'imputation ont été envisagées pour imputer les blocs de données peu fiables. Les approches locales (KNN, lissages) ne sont pas directement applicables pour estimer des séquences longues de données manquantes, liées par exemple à un dysfonctionnement de la cellule. Il est possible d'utiliser des statistiques saisonnières pour combler ces données manquantes, toutefois cette approche ne prendra pas en compte les spécificités de l'année en cours, et ne tirera pas profit des autres données. Des imputations globales, basées sur les corrélations inter-séries, via des modèles de régression, en utilisant par exemple avec des algorithmes tels que NIPALS (Nonlinear Iterative Partial Least Squares) ou MICE (Multiple Imputation by Chained Equations), ou d'autres approches basées par exemple sur les forêts aléatoires ou les réseaux de neurones, semblent préférables.

Plusieurs de ces algorithmes ont été testés (cross validation). Ici, le choix s'est porté sur un algorithme de random forests pour estimer les séquences manquantes. Plusieurs modèles avec différents ensembles de prédicteurs ont été testés et évalués par validation interne et externe (non développé ici).

Les données ont été ici imputées en utilisant l'algorithme missForest sur un ensemble de prédicteurs incluant les données du Panoramique des Dôme, des données météorologiques et des données calendaires, à l'exclusion des autres données d'éco-compteurs, en raison du manque de fiabilité ou de représentativité de la plupart des séries d'éco-compteur, même si certaines auraient pu être éventuellement intégrées.

On peut voir sur Figure 38, les séries 2019 des flux entrants (IN) sans les pics (en bleu), avec les valeurs imputées (en rouge).

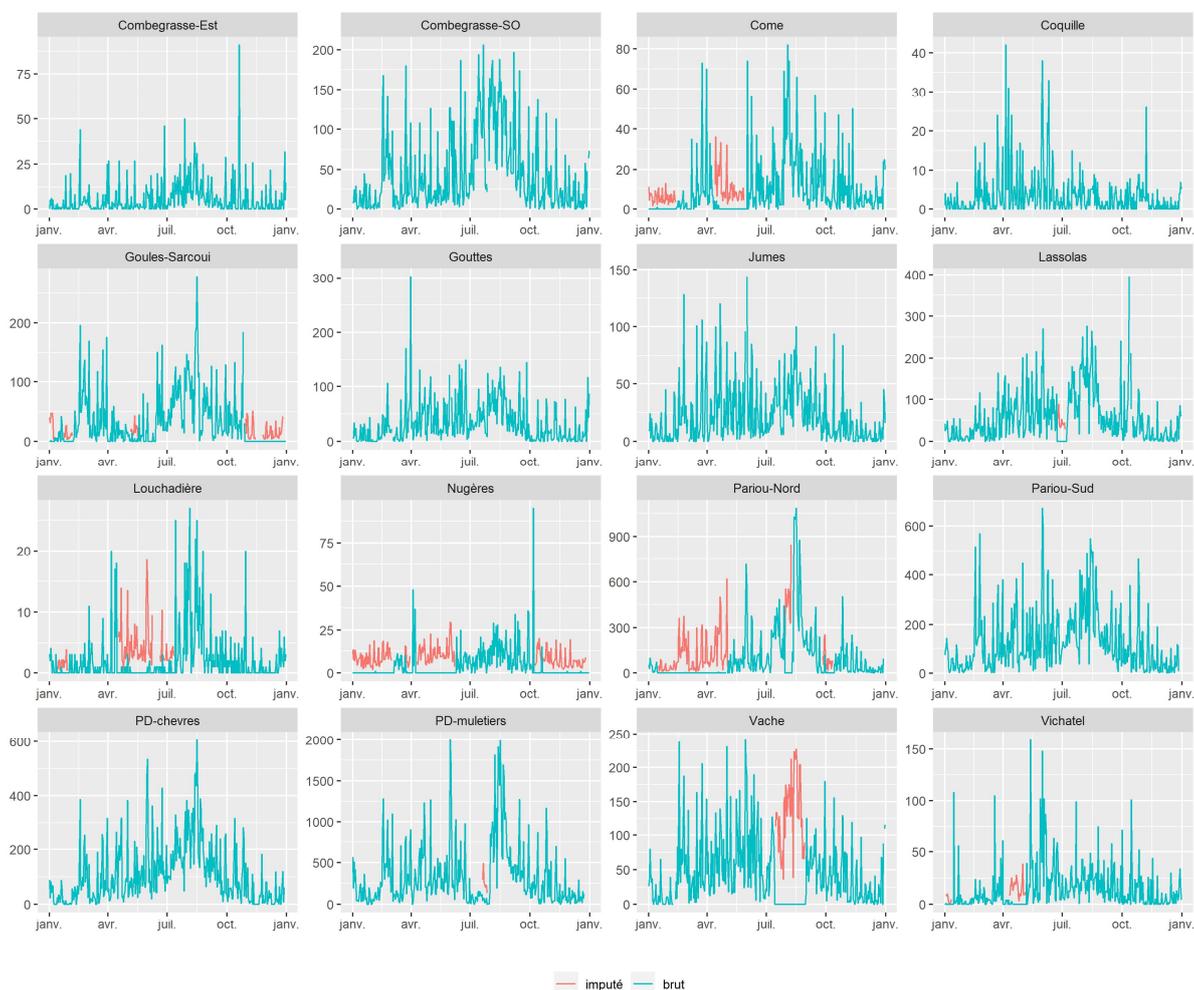


Figure 38 - Séries brutes (bleu) et imputées (rouge)

### 3.1.3. COMBINAISON ET EXTRAPOLATION

Pour minimiser les problèmes de doubles comptes, notamment sur la zone Puy de Dôme-Puy de Pariou, un cumul par zone été réalisé, c'est-à-dire en considérant l'ensemble des entrées/sorties sur la zone.

Puis, les données ont été extrapolées, en utilisant un coefficient multiplicatif établi à partir des informations fournies par le PNR en février 2021, pour tenir compte des pertes (sentiers alternatifs...) et des autres sentiers périphériques non-équipés (ex. tour du Puy de Dôme). Ces coefficients doivent être vus comme des paramètres révisables permettant d'obtenir une estimation plus proche de la fréquentation réelle des circuits de la CDPFL, sachant que d'autres circuits non-équipés ne sont pas considérés (ex. Montagne de la Serre, Faille de Limagne).

---

## COMBINAISON PAR SITE

Pour estimer le nombre de visites sur les différents sites de la Chaîne des Puys, les passages ont été agrégés en calculant la moyenne des flux journaliers entrant et sortant, pour chacun des éco-compteurs. Ainsi, par exemple, pour le puy de Louchadière, le nombre de passages est calculé ainsi :

$$Louchadière = (Louchadière_{IN} + Louchadière_{OUT})/2$$

De même, pour les trois sites suivants, équipés de deux éco-compteurs : Combegrasse, Jumes-Coquille, et Vache-Lassolas, là encore, le nombre de passages a été calculé à partir de la moyenne des flux entrant et sortant sur les deux éco-compteurs, soit respectivement :

$$Combegrasse = (Combegrasse.Est_{IN} + Combegrasse.SO_{IN} + Combegrasse.Est_{OUT} + Combegrasse.SO_{OUT})/2$$

En revanche, pour le Puy de Dôme (PD) et le puy de Pariou, afin d'éviter les doubles comptes (personnes visitant les deux puys dans la même journée), l'agrégation a été faite ainsi :

$$PD = PD.chevres_{IN} + PD.muletiers_{IN}$$

Pour le Puy de Dôme, sont également prises en compte, par ailleurs, les montées par le Panoramique des Dômes.

Au vu des courbes de fréquentation horaire, il a été considéré qu'une part importante des combinaisons Puy de Dôme – Pariou se faisait en passant par les éco-compteurs du sentier des chèvres et celui de Pariou-Sud, et que l'essentiel des entrées depuis le parking des Goules vers le Pariou se faisait via l'éco-compteur de Pariou-Nord. C'est pourquoi, seul les flux entrants via Pariou-Nord sont pris en compte pour les visites depuis le parking des Goules.

Enfin, le calcul repose également sur les deux hypothèses suivantes :

- Le puy de Côme, interdit et isolé, se fait plutôt hors circuit combiné ;
- Les autres puys (Cliersou, Grand et Petit Suchet) ne sont pas visités seul (hors Pariou ou Puy de Dôme).

Restent la possibilité d'autres accès par la Fontaine des bergers ou par Vulcania, ainsi que d'autres circuits, également très fréquentés, ne passant pas par les puys. C'est pourquoi il a été proposé d'extrapoler ces comptages, pour essayer de se rapprocher de la fréquentation réelle du site.

---

## EXTRAPOLATIONS

Ayant sollicité l'expertise des gardes nature du PNR des Volcans d'Auvergne en février 2021, pour avoir une idée des écarts possibles entre les mesures des éco-compteurs et la fréquentation réelle des sites, nous avons utilisé leurs premières estimations pour définir des coefficients (Tableau 5), pour essayer de se rapprocher de la fréquentation réelle, notamment dans l'optique de comparer ces estimations de fréquentation aux données de Flux Vision.

Nous avons donc extrapolé les nombres de visites sur les différents sites, en multipliant les estimations de chaque site par le coefficient correspondant. Ces coefficients sont sans doute discutables, et sont évidemment révisables. En outre, cette première correction repose sur l'hypothèse simple mais probablement fautive d'une proportionnalité des écarts quelle que soit la saison, et devrait être ajustée plus finement, en fonction des connaissances qui pourraient être acquises en étudiant la circulation réelle sur et autour des puys.

## 2.3. AUTRES SOURCES

---

### PANORAMIQUE

Pour le Panoramique, des séries journalières (montée / descentes) ont simplement été créées à partir des données par train (aller/retour), données qui semblent a priori complètes et sans anomalies apparentes.

---

### BILLETTERIES

Pour les différents sites payants, nous avons obtenu par le CD63, en début de projet, les données annuelles (cumul annuel des entrées) jusqu'en 2017. Puis, les entrées annuelles 2018 et 2019 nous ont été transmises le 23 mars 2021.

Le 29 novembre 2019, nous avons fait une demande de données plus précises sur le mois d'août 2019 (données journalières set, si possible, origine des visiteurs), auprès du Volcan de Lemptégy et de Vulcania, afin de pouvoir tester la fiabilité des données de Flux Vision sur la sous-zone centrée sur les deux sites.

Le 12 décembre 2019, Lemptégy nous a transmis les données journalières sur août 2019, accompagnées de la ventilation par origine (pays, département), pour chacune des dates. De plus, Lemptégy nous a transmis le 8 avril 2020 des données mensuelles, depuis janvier 2018 jusqu'à mars 2020. Enfin suite à une erreur signalée en janvier 2022 dans le tableau des données par origine, un correctif nous a été envoyé le jour même.

En revanche, Vulcania n'a pas souhaité nous transmettre ces données sur le mois d'août, malgré notre engagement à ne pas publier les données, et à n'en faire qu'un usage interne, à la seule fin de tester la fiabilité de Flux Vision. Une lettre du conseil départemental a également été adressée à Vulcania pour appuyer notre demande, mais sans succès. Nous avons relancé Vulcania en 2020, en leur proposant de ne nous fournir que des données d'entrées pour quelques dates en août 2019. Finalement, après plusieurs relances, nous avons obtenu, le 17 mars 2021, un échantillon de 8 valeurs, correspondant aux entrées journalières sur 8 dates du mois d'août 2019.

A partir des données disponibles sur les sites payants, (cumuls annuels) et des données récupérées (données mensuelles de Lemptégy, calendriers d'ouverture), les données

journalières ont été estimées pour chaque site. Les données mensuelles ont été estimées à partir de la répartition par mois des entrées sur le Volcan de Lemptégy. Les cumuls mensuels ont ensuite été ventilés sur les jours d'ouverture en ajustant la répartition selon l'importance relative des entrées du Panoramique des Dômes sur la série des jours d'ouverture. Enfin, les entrées journalières connues, sur le mois d'août 2019, à Lemptégy et à Vulcania, ont été réinjectées dans les séries respectives.

A noter que sur Volvic une seule série a été reconstituée en considérant les jours d'ouverture de la Maison de la Pierre (plus étalés que celle de l'Espace Information des Eaux de Volvic) et le cumul annuel de l'Espace Info (car plus élevé que celui de la Maison de la Pierre).

$$N_j = I_{ouv.}^j \times N_{tot} \times \frac{N_{Lemptégy}^{mois\ k}}{N_{Lemptégy}^{annuel}} \times \frac{N_{PDD}^j}{\sum_{j \in mois\ k} (N_{PDD}^j \times I_{ouv.}^j)}$$

$N_j$  est le nombre de visites à la date  $j$

$I_{ouv.}^j$  est une indicatrice de l'ouverture du site

$N_{tot}$  est le nombre total de visites sur l'année

$N_{Lemptégy}^{mois\ k}$  est le nombre de visites sur le mois  $k$

$N_{PDD}^j$  est le nombre de montées au Panoramique à la date  $j$

---

## AYDAT

L'absence de suivi de la fréquentation sur Aydat, ainsi que l'absence de chiffres de fréquentation permettant d'avoir un ordre de grandeur de cette fréquentation, ne permettaient pas de faire une estimation du nombre de visites sur ce site.

Toutefois, suite à un échange avec le service tourisme de Mond'Arverne, en janvier 2021, les données journalières d'entrée dans l'office de tourisme d'Aydat, situé à proximité de la plage, nous ont été transmises. Au cours de cet échange, nous a également été communiquée une estimation approximative du pic de fréquentation durant l'été, c'est-à-dire une estimation du nombre maximal de personnes présentes – au même moment - sur la plage, de l'ordre de 2 000 personnes. Une première tentative d'estimation de la fréquentation d'Aydat a donc été entreprise, à partir de ces données.

En faisant l'hypothèse - très discutable, au moins en basse saison - que la fréquentation de l'office de tourisme est bien corrélée avec la fréquentation globale du site, une estimation a été réalisée sur la base des données journalières d'entrée, seul proxy disponible de la fréquentation totale, ajustée sur le pic de fréquentation journalière ( $N_{max}$ ), c'est-à-dire en utilisant le ratio des fréquentations maximales de l'office de tourisme et de la plage.

$$V^j = n^j \times \frac{N_{max}}{\max(n^j)}$$

Puisque l'on ne dispose que d'une estimation instantanée de ce pic, il est nécessaire d'estimer la fréquentation journalière correspondante. Pour ce faire, en estimant une distribution quasi-gaussienne des arrivées ( $A_i$ ) entre 8h et 18h, avec une hypothèse de durée moyenne de présence ( $d$ ) de 4h, on peut, sachant que le pic est de l'ordre de 2500 personnes, estimer grossièrement, par simulation, la fréquentation journalière maximale, qui serait, sous ces hypothèses, de l'ordre de 4000 personnes.

$$N_{max} = \operatorname{argmin}(\max(N_t) - N_0)$$

$$\text{avec } N_t = \sum_{i=0}^n A_i \in [t - d; t + d]$$

$$\text{et } d = 4 \text{ et } A_i \sim \mathcal{N}(13, 2)$$

Ayant obtenu la série de visites journalières ajustée sur ce pic, il reste à imputer les données manquantes correspondant aux dates et périodes de fermeture de l'office du tourisme. Ces portions manquantes de la série ont été imputées à partir d'un modèle de régression des visites sur Aydat sur le nombre de montées en Panoramique (Figure 39).

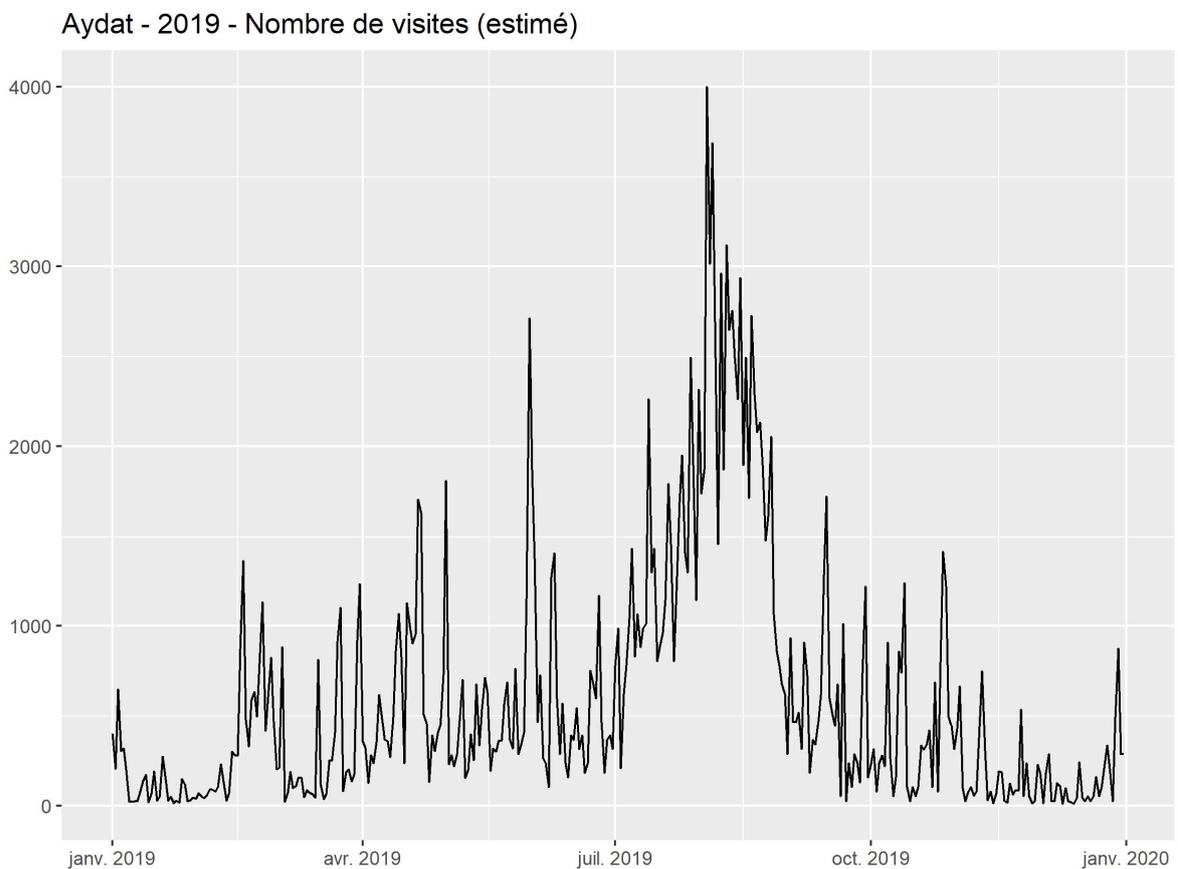


Figure 39 - Fréquentation journalière d'Aydat (estimée)

## 2.4. ESTIMATION GLOBALE

### 2.4.1. AGREGATION DES SOURCES

A partir des séries reconstituées pour les différents sites, un simple cumul de ces séries a permis de reconstituer la série de la fréquentation totale de la zone CDPFL (Figure 40). Cette agrégation simple fait l'hypothèse que, pour les sites considérés, un même visiteur ne visitera pas plusieurs sites dans une même journée, ou que le nombre de visiteurs se rendant sur plusieurs sites dans une même journée (par ex. au Puy de Dôme et à Vulcania) est négligeable.

Cette hypothèse est très discutable (et probablement fausse), mais elle pourra être réévaluée en fonction de la disponibilité d'informations fiables sur les visites multiples. Les données d'enquête pourraient être utilisées à cet effet, mais ces informations sur le programme réel de visite des personnes interrogées, qui sont du déclaratif, et souvent même du prospectif (visites prévues), et surévaluent probablement le nombre de visites. Une estimation du nombre de visites sur un sous-échantillon plus fiable a été utilisée pour l'estimation globale dans (Marsat, 2022). On pourrait également considérer les données de durée de présence de Flux Vision, en considérant que la probabilité de faire plusieurs visites dans une même journée pourrait être estimée par une fonction de la durée de présence.

En outre, on peut penser que, pour des sites payants et des sites de randonnée, impliquant des temps de présence relativement longs, l'hypothèse peut se justifier quelque peu, sauf peut-être en période estivale, où les journées sont plus longues et permettent plus facilement de visiter plusieurs sites dans la même journée.

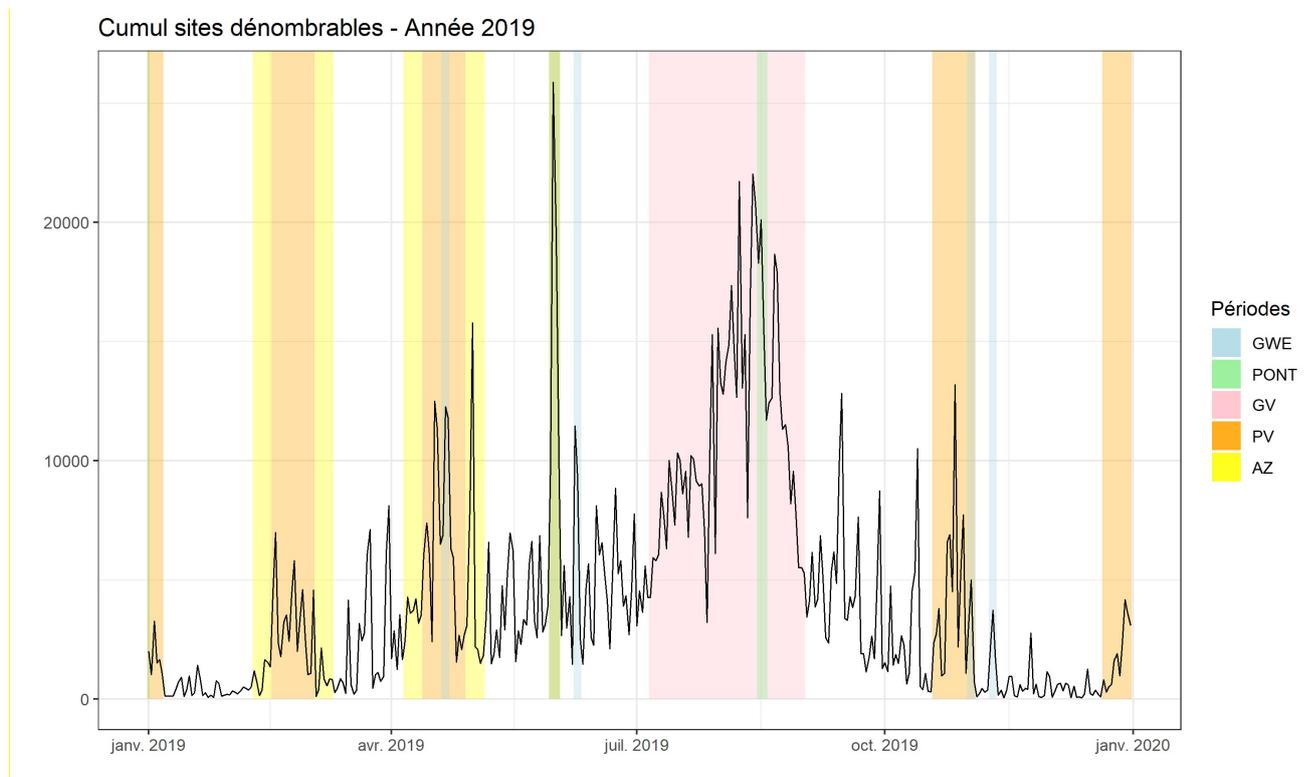


Figure 40 - Série complète de la fréquentation de la CDPFL en 2019 obtenue par agrégation des séries des sites dénombrables (GWE = week-ends prolongés, GV = grandes vacances, PV = petites vacances, AZ = autres zones)

---

#### 2.4.2. CONVERSION DES VISITES EN VISITEURS

Obtenir un nombre de visiteurs à partir d'un nombre de visites suppose de connaître les composantes de la fréquentation et le nombre de visites annuel pour chacune d'elles, tout en sachant que la composition de la fréquentation et les nombres de visites correspondants fluctuent fréquemment au cours de l'année.

A minima, il est nécessaire de connaître les **proportions d'excursionnistes et de touristes**, ainsi que le nombre moyen de visites pour chacune de ces fractions.

La part des touristes sur l'ensemble des visiteurs pourrait être évaluée à partir de la ventilation par origine des entrées du Panoramique des Dômes, mais on peut s'attendre à une part quelque peu surestimée de touristes. Des données d'enquête peuvent également être utilisées, à condition d'avoir un échantillon représentatif et/ou un redressement fiable. Enfin, il serait possible a priori d'utiliser les données Flux Vision, si elles semblent fiables, pour connaître cette composition.

Par ailleurs, il faut également connaître le **nombre de visites dans l'année**, c'est-à-dire le nombre de visites durant le ou les séjour(s) pour les touristes et la fréquence des visites pour les excursionnistes. Ici, a priori, seules les données d'enquête peuvent permettre d'estimer ces paramètres.

Sur la base des données de l'enquête 2019-2020, on obtient pour les 1.65 millions de visites estimées en 2019 à partir des données agrégées des sites dénombrables :

- 560 014 touristes,
- 222 466 excursionnistes.

## 3. METHODES ALTERNATIVES

### 3.1. COMPTAGES ROUTIERS

#### 3.1.1. EXPERIMENTATION

---

##### CONTEXTE

Une première difficulté provient, comme on l'a vu (cf. 1.6), de l'ouverture du site et de la présence de nombreux axes routiers importants, avec des mobilités résidentielles importantes ainsi que du transit, notamment entre autoroutes. Or, globalement, ces passages ne concernent pas la fréquentation de tourisme et de loisirs.

Par ailleurs, de nombreux parking permettent d'accéder aux sites de la zone CDPFL : des parkings officiels, mais également des parkings non officiels et sauvages à divers endroits de la zone (Tableau 6), ainsi qu'un important stationnement le long des routes en période estivale, lorsque les parkings officiels sont pleins.

Enfin, si l'on considère la dizaine de parkings permettant d'accéder aux circuits de randonnée ou aux diverses attractions du site (Figure 41), aucun n'est situé sur une extrémité de route sans issue, comme c'est le cas dans de nombreux sites naturels, en particulier dans les massifs, où il est possible de compter les véhicules des visiteurs, de manière exhaustive et spécifique (Figure 42).

Toutes ces caractéristiques rendent difficile un suivi de la fréquentation fondé uniquement sur des comptages de véhicules.



**1. Parking Beauregard bourg (commune de Saint-Ours-les-Roches) :** le long de la D576, un stationnement permet d'accéder à 2 PR (Jumes Coquille et Louchadière) ainsi qu'à la sculpture de Thierry Courtadon. Sa capacité permet d'accueillir 35 places pour des véhicules légers + 2 places pour les personnes à mobilité réduite. Le stationnement pour les cars est aussi possible (2).

**2. Parking Beauregard bourg (commune de Saint-Ours-les-Roches) :** NE PLUS INDIQUER CE STATIONNEMENT.

**3. Parking des puys des Gouttes et Chopine (commune de Saint-Ours-les-Roches) :** 50 places pour des véhicules légers stationnés en épis, stationnement possible pour 5 à 6 bus.

**4. La Fontaine du Berger (commune d'Orcines) :** tolérance sur la partie haute du terrain ; par contre, interdiction de stationnement sur le terrain en aval à l'intérieur du camp militaire (panneau d'indication) – stationnement possible pour 20 véhicules légers. Interdiction faite au bus avec toutefois la possibilité d'effectuer des manœuvres de retournement à l'intérieur du terrain militaire.

**5. Parking du puy de Dôme (commune d'Orcines) :** 1000 places de stationnement pour les véhicules légers, 13 places de stationnements pour les bus et 10 places de stationnements pour les camping-cars uniquement en journée.

**6. La Font de l'Arbre (commune d'Orcines) – Parking vers le Village Auvergnat :** dépose minute pour les bus. À partir des parkings de la Font de l'Arbre, suivre les cheminements organisés pour rejoindre le site du puy de Dôme **GRAND SITE DE FRANCE** et les points de départ des PR.

**7. Parking col de Ceysat (commune de Ceysat) :** 110 places pour les véhicules légers et 2 places pour les camping-cars uniquement en journée. Pas de possibilité de stationnement pour les bus, uniquement de la dépose minute.

**8. Parking des puys de la Vache et de Lassolas (commune de Saint-Genès-Champanelle) :** 50 à 60 places pour les véhicules légers, stationnement possible pour 5 bus.

**9. Parking Maison du Parc à Montlosier (commune d'Aydat) :** 30 places pour les véhicules légers et 3 places bus.



**10. Nouveau Parking Combegrasse (commune d'Aydat) :** 30 places de véhicules légers et 2 bus.

**11. Parking des Goules (commune d'Orcines) :** 145 places de stationnement pour les véhicules légers, 7 places pour les bus et 4 à 5 places pour les camping-cars uniquement en journée. Une trentaine de places sont accessibles en hiver.



**12. Parking du col de la Ventouse :** l'accès à l'aire du col de la Ventouse le long de la D2089 va faire l'objet de travaux à partir du mois de mai pour une durée de 2 mois afin que l'aire devienne accessible aux randonneurs en provenance depuis Clermont-Ferrand. La capacité du stationnement reste inchangée : 16 places pour les véhicules légers et 5 places pour les cars. Toilettes ouvertes d'avril à fin octobre.

Tableau 6 - Liste des parkings de la zone d'étude (CD63, 2019)

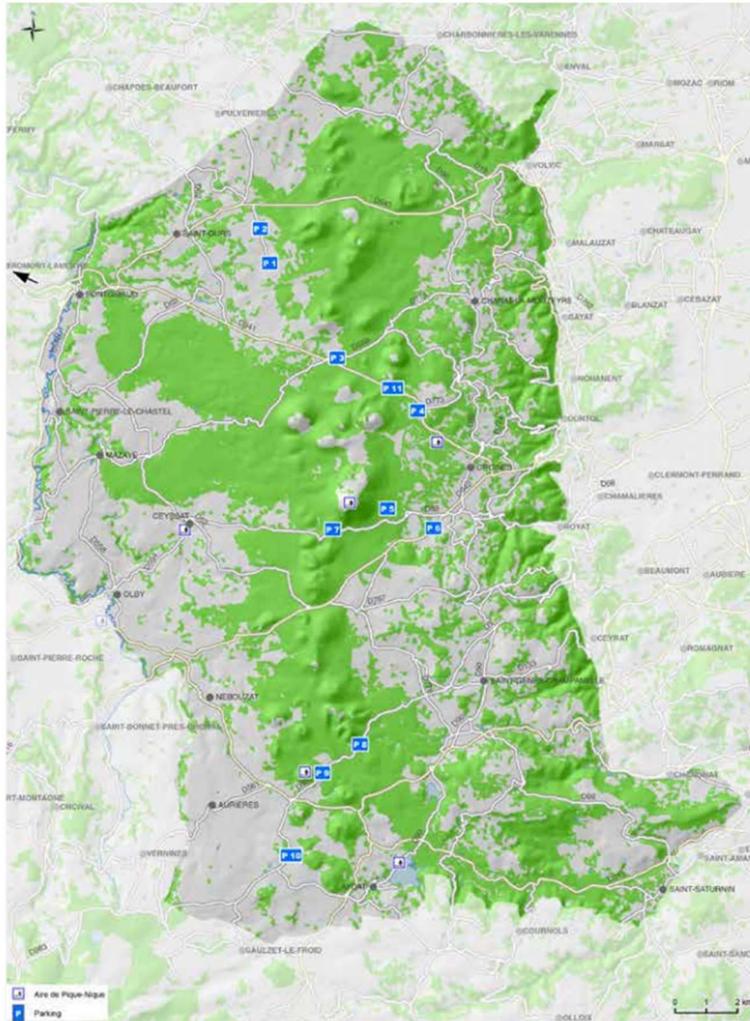


Figure 41 - Carte des parkings de la zone d'étude (CD63, 2019)

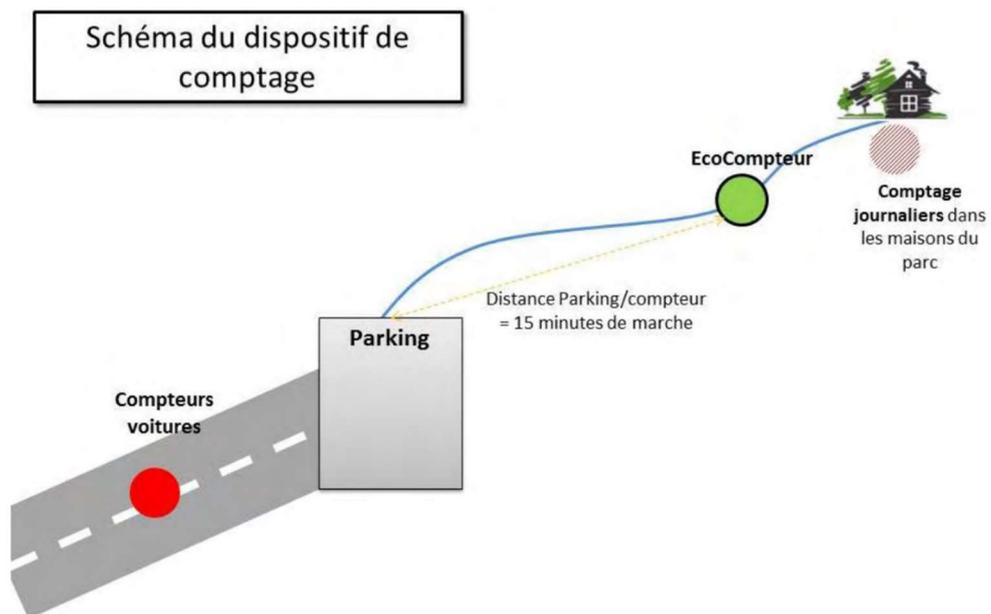


Figure 42 - Dispositif de comptage mis en place dans le Parc National des Écrins sur un parking situé en bout de route (source : (PN des Écrins, 2011))

---

## MISE EN PLACE

Suite à une réunion avec le Service des Routes organisée début 2019 par le CD63, au cours de laquelle nous ont été présentés les différents dispositifs de comptage routier, il a été proposé de mettre en place une expérimentation au cours de l'été, période pendant laquelle les compteurs sont très peu utilisés.

Initialement, nous souhaitions voir avec le Service des Routes s'il existait des dispositifs de comptage sur les parkings permettant de compter les véhicules stationnés ou les véhicules entrant et sortant du parking. Nous avons été informés que le Service des Routes ne disposait pas de système de comptage sur parking, mais il nous a été signalé par la suite qu'un test ponctuel de comptage vidéo serait réalisé durant l'été sur le parking du Volcan de Lemptégy.

Nous avons demandé si des comptages classiques pouvaient être réalisés en amont et en aval des parkings du site, afin d'essayer d'estimer le nombre de véhicules entrant et sortant du parking. En effet, il semblait intéressant d'essayer d'estimer ces flux à partir de compteurs placés en amont et en aval des parkings, en considérant que le pas de temps de 6 minutes, sur lequel sont agrégées les données, était suffisamment court pour estimer le nombre de véhicules entrant et sortant du parking sur ces intervalles.

Le Service des routes nous a cependant bien signalé que ces dispositifs mobiles de comptage perdaient rapidement en fiabilité au fil des jours et, qu'au bout d'une semaine, le tube disposé en travers de la route était tellement écrasé qu'il ne permettait plus d'obtenir une mesure fiable du nombre de passages.

Nous avons eu ensuite des échanges répétés avec Willy Dehayes, au cours du mois de juin 2019, afin de mettre au point le protocole. Nous avons proposé de placer 10 compteurs pour couvrir les principaux parkings du site (Figure 43), avec deux poses de compteurs autour du 14 juillet et autour du 15 août. Le Service des Routes nous a proposé de poser les compteurs du 8-9 juillet jusqu'au 18 juillet pour la première période, et du 12 août jusqu'au 22 août pour la seconde période.

Les compteurs devaient être placés ainsi :

- Deux compteurs autour du parking des Goules sur la D941 ;
- Un compteur à l'est de Lemptégy pour permettre, avec le compteur semi-permanent déjà en place, de suivre les entrées et sorties du parking de Lemptégy, voire également, avec le compteur à l'ouest du parking des Goules, d'estimer le stationnement sur le parking de Vulcania, à condition de négliger le trafic des routes secondaires vers Ceysnat et Chant-la-Mouteyre, au rond-point de Vulcania (hypothèse à vérifier) ;
- Trois compteurs sur la route du col de Ceysnat permettant d'avoir l'amont et l'aval du grand parking du Panoramique des Dômes, ainsi que l'amont et l'aval du petit parking du col de Ceysnat, point de départ du sentier des muletiers ;
- Deux compteurs à l'amont et à l'aval du parking des puys de Vache et Lassolas,
- Et, enfin, deux autres compteurs, sur la D90, aux deux extrémités de la plage d'Aydat, complétant le dispositif.

## TEST DE FIABILITE

Pour évaluer la fiabilité des comptages routiers, nous avons prévu des sorties sur le terrain pour compter de visu les véhicules pendant environ une heure sur différents parkings (Figure 44). Un protocole et une grille de saisie ont été conçus pour bien relever les nombres et types de véhicules pour chacun des différents flux (Figure 45). Les relevés ont été réalisés les 11 et 16 juillet pour la première période de pose, et les 13 et 23 août pour la seconde période. Nous avons également profité de ces sorties pour vérifier si les flux via les routes secondaires paraissaient - ou non - négligeables, et pour estimer grossièrement le nombre moyen de personnes par véhicule, ainsi que la variabilité de ce nombre. Le choix de ces dates de fortes fréquentation visait à pouvoir observer des comportements particuliers ou imprévus, ainsi qu'à comptabiliser l'importance du stationnement hors parking le long des routes.

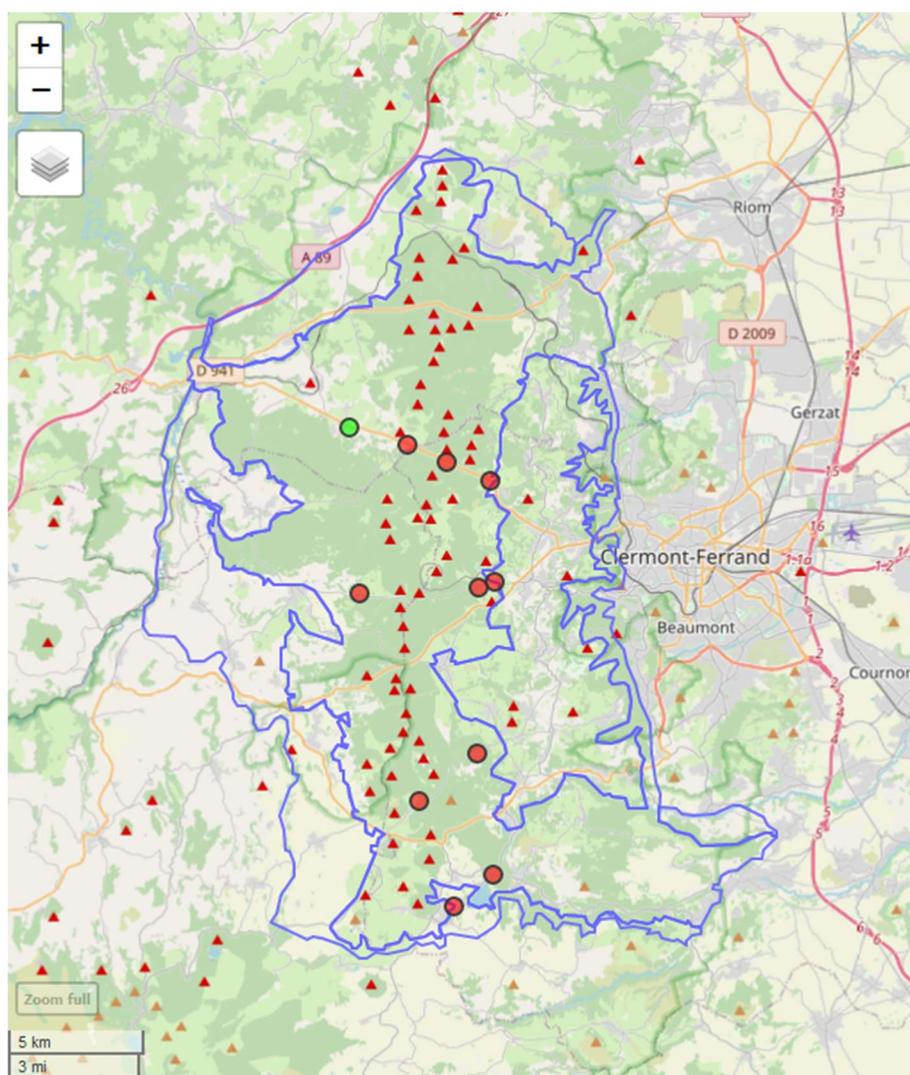


Figure 43 - Carte de la répartition des compteurs routiers posé durant l'été 2019 (en vert : compteur semi-permanent déjà en place)

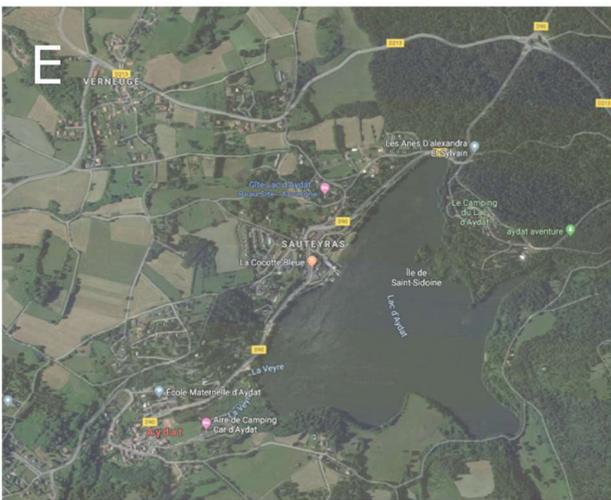


Figure 44 - Photos des lieux de comptages (A = rond-point de Vulcania, B = parking des Goules, C = parking du col de Ceysnat, D = parking du Panoramique, E = lac d'Aydat, F = parking de Vache-Lassolas)

DONNEES COLLECTEES

Les données fournies par le Service des routes à l'issue de l'expérimentation présentent un certain nombre de problèmes :

- Certains compteurs n'ont pas fonctionné : compteur Nord à Aydat (Figure 45-D) et compteur Est du parking des Goules, pour lesquels les données sont indisponibles) ;
- Certains ont très vite dysfonctionné (compteur Sud du parking de Vache-Lassolas, où seule la première semaine est exploitable) (Figure 45-C et Figure 46) ;
- D'autres compteurs présentent plus ou moins d'incohérence, lorsque l'on regarde les flux entrant et sortant, cumulés date par date, ceux-ci devant être identiques ou proches (par ex. à Lemptégy, Figure 47, écarts parfois importants, allant jusqu'à 1 000 voire 1 500 véhicules par jour) ;
- Enfin, et surtout, les données ne sont disponibles finalement qu'à un pas de temps de 1h, et non de 6 minutes comme nous l'avions compris, le pas de 6 minutes étant « uniquement disponible avec la station SIREDO, station fixe avec relevé à distance », comme nous l'a confirmé Willy Deshayes.

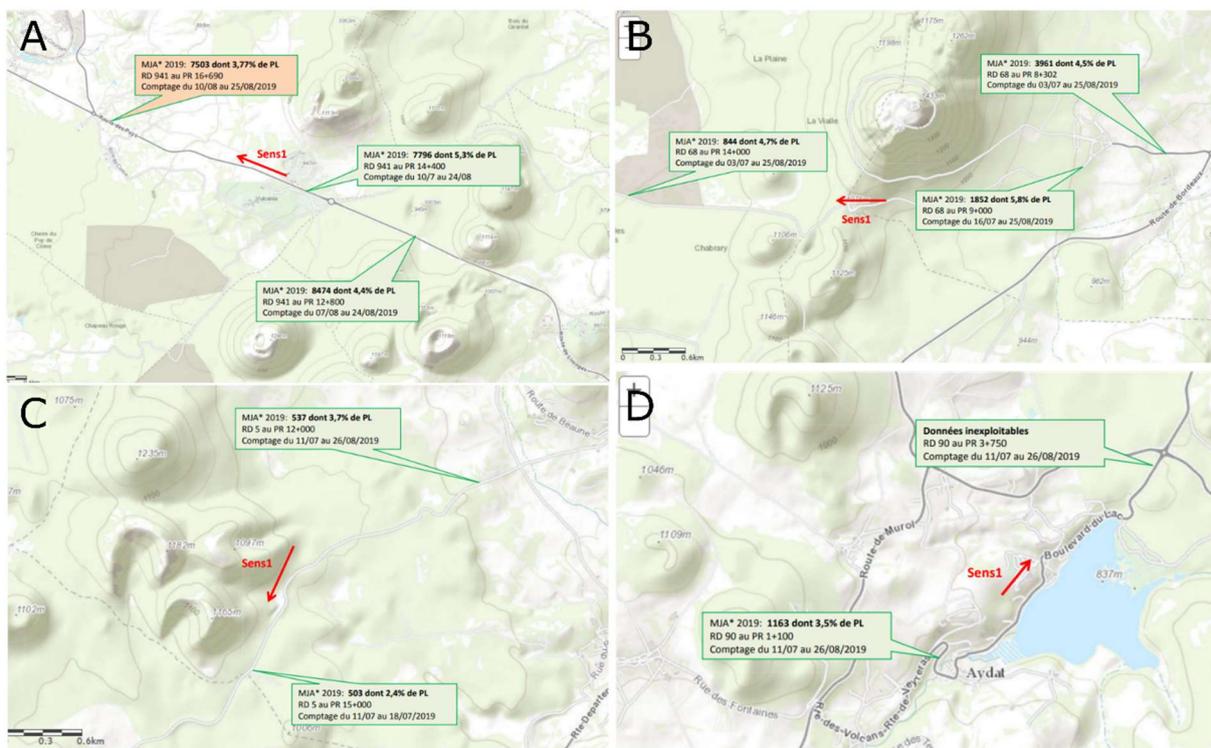


Figure 45 - Comptages réalisés par le Service des routes du CD63

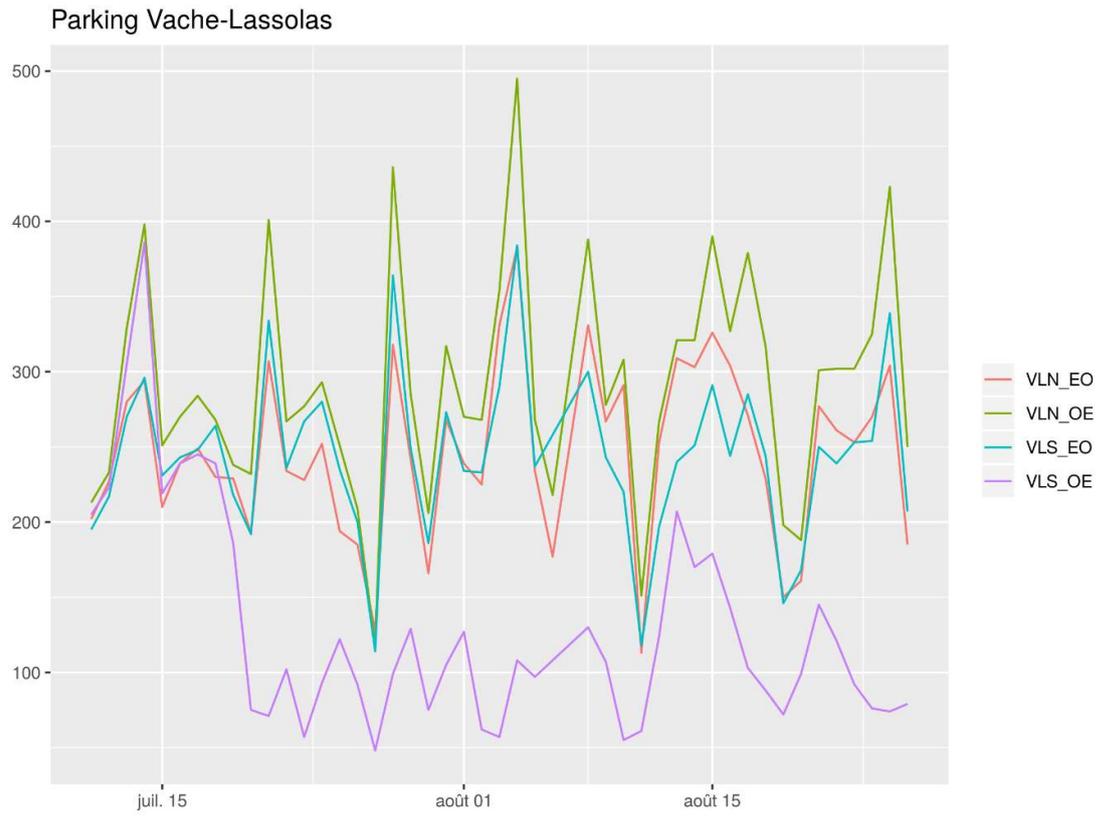


Figure 46 - Comptages routiers en amont et aval du parking de Vache-Lassolas, avec dysfonctionnement du compteur Sud après le 18 juillet 2019

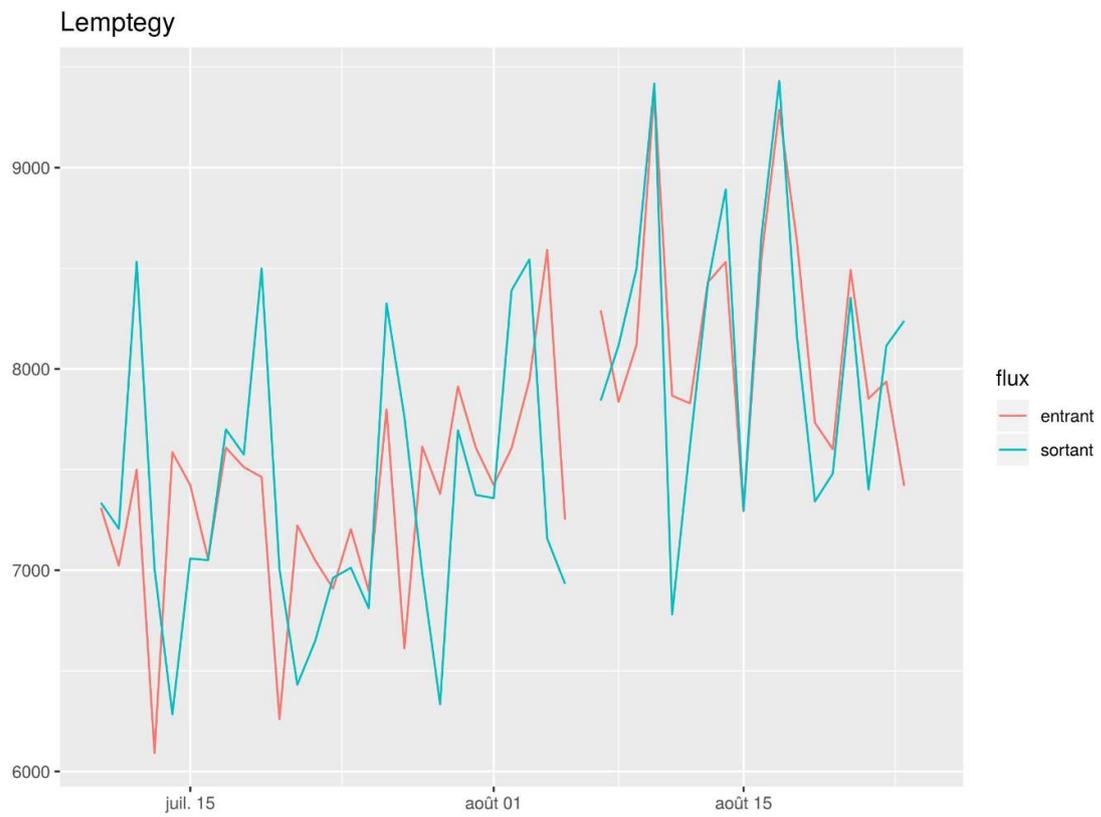


Figure 47 - Flux entrant et sortant pour les compteurs amont et aval au niveau du parking de Lemptegy

On peut regarder, sur des séries présentant moins de problèmes, telles que celle du parking du Panoramique des Dômes, la cohérence, pour une date donnée, entre les entrées-sorties du parking et les passages comptabilisés par les éco-compteurs et les allers et retours par le train (Figure 48).

En première analyse, il est possible ici, en négligeant les flux venant de Ceysnat, de voir la concordance des entrées du parking avec les montées (avec un retard attendu des montées) et celle des sorties du parking avec les descentes (avec un retard attendu des sorties de véhicule). En comptant une moyenne de 3 personnes par véhicule, on obtient, en convertissant les entrées-sorties du parking en nombres de personnes, des courbes quasiment identiques dans chaque sens, avec un léger décalage correspondant possiblement au délai moyen entre le passage de l'éco-compteur ou la montée dans le train et l'arrivée ou le départ du parking.

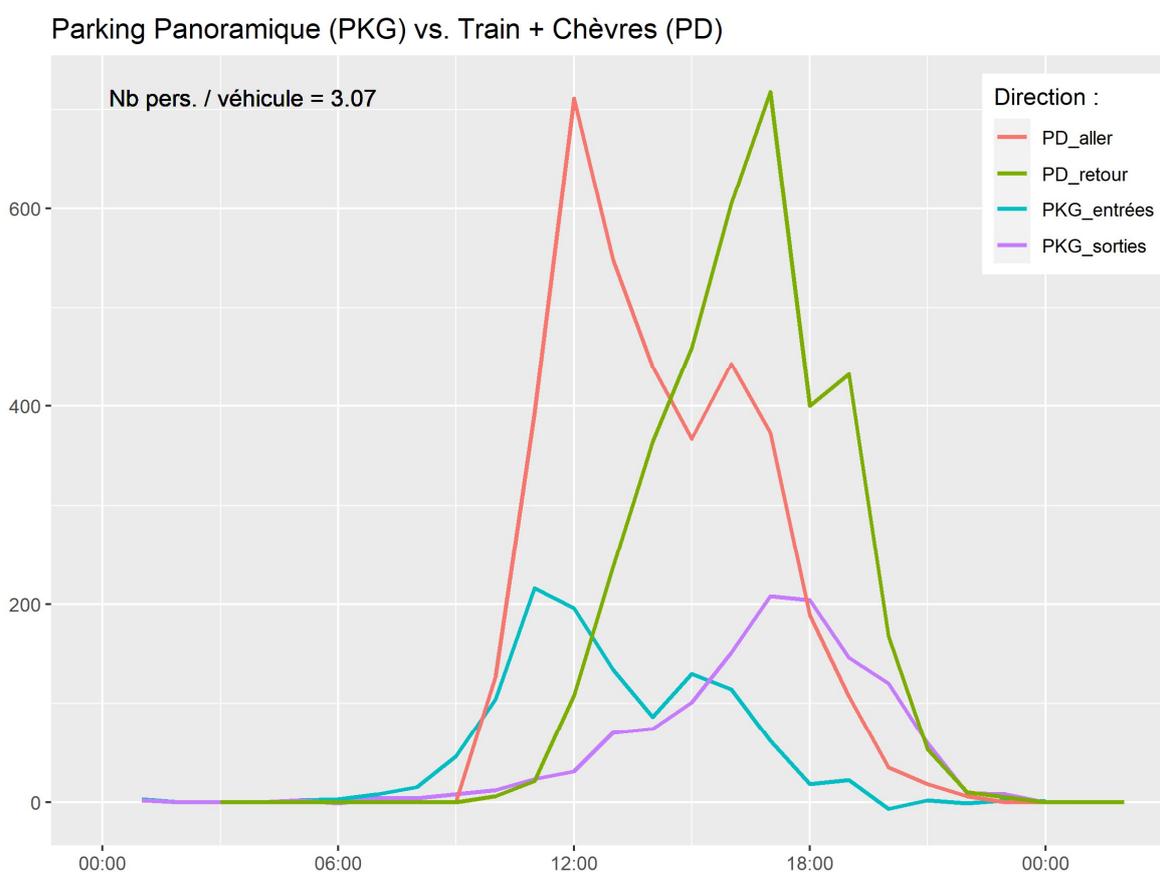


Figure 48 - Comptage des véhicules sur le parking du Panoramique vs. cumul train + sentier des chèvres

## ESTIMATION

Le principe était, pour un parking donné, de placer deux compteurs, respectivement en amont et en aval du parking (Figure 49), afin de mesurer les flux traversant au niveau du carrefour, et d'essayer d'estimer les flux entrant et sortant vers et depuis le parking (Figure 50). A partir d'un modèle probabiliste, on peut essayer d'estimer les flux entrant ( $E = E1 + E2$ ) et sortant ( $S = S1 + S2$ ), à partir des flux traversant mesurés, et d'en déduire le nombre de véhicules venus sur le parking.

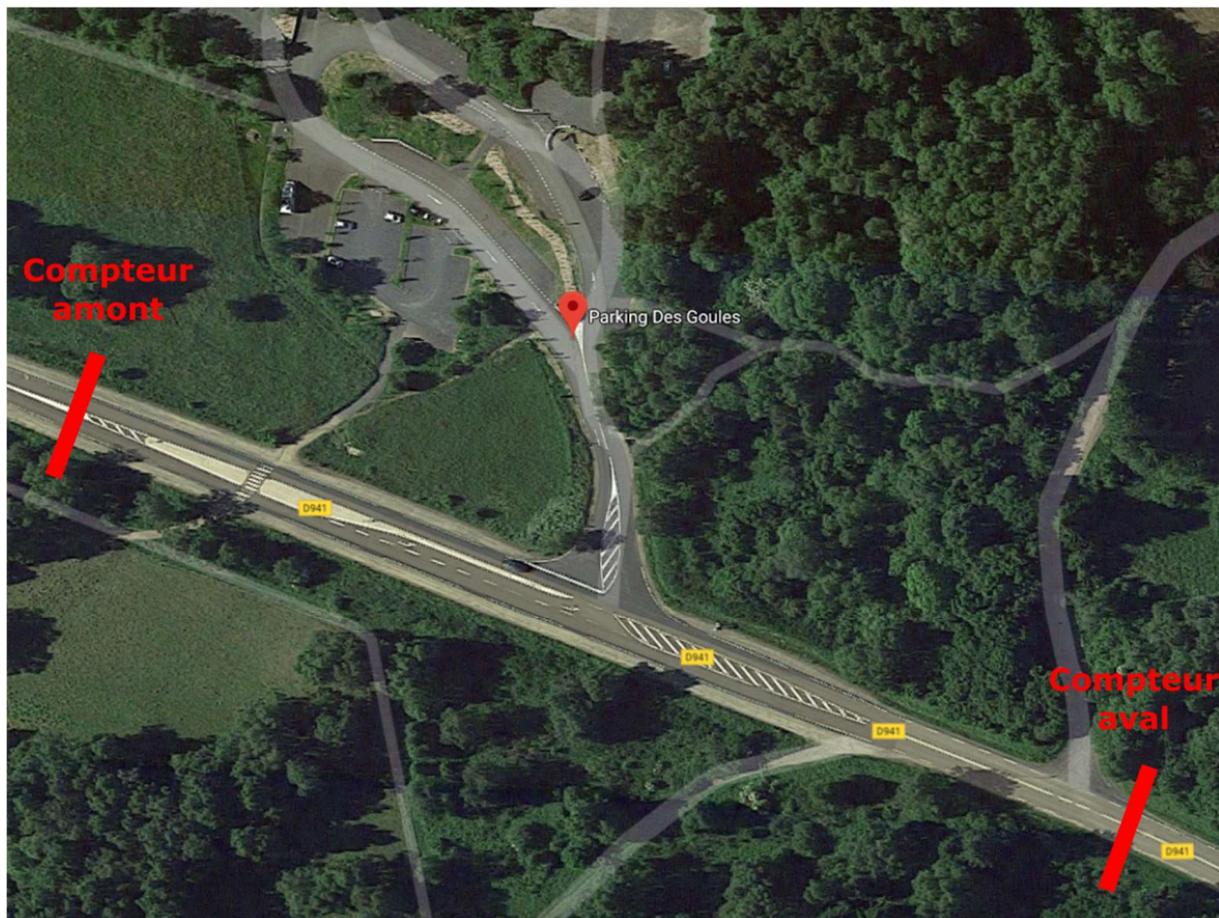


Figure 49 - Photo du parking des Goules avec représentation schématique du positionnement des compteurs routiers en amont et en aval

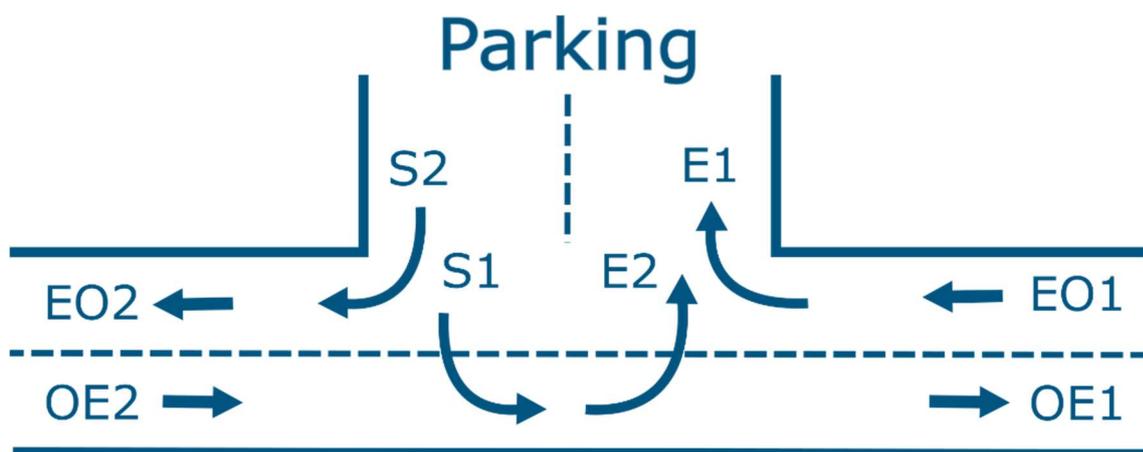


Figure 50 - Schéma des principaux flux au niveau du parking

Si l'on confronte ces données (nombre de véhicules en stationnement) avec le nombre de personnes comptées sur le site (via les passages par des éco-compteurs ou le nombre d'entrées d'un site payant), on peut vérifier la corrélation des séries, dans l'hypothèse où le facteur multiplicatif est constant, c'est-à-dire si le nombre moyen de personnes par véhicule est constant, et si les autres modes de transport évoluent proportionnellement.

On peut voir sur Figure 51, les courbes des montées par le sentier des muletiers et des nombres estimés de véhicules ayant stationné sur le parking du col de Ceysat (multipliés par 5). Ces courbes sont très corrélées à partir du 1<sup>er</sup> août 2019 et très proches, suggérant qu'à un facteur multiplicatif de 5 près, on peut déduire une courbe à partir de l'autre. Ce n'est pas le cas avant le 1<sup>er</sup> août, mais les données d'éco-compteur ne sont clairement pas fiables sur le mois de juillet (Figure 52). Ainsi pourrait-on envisager, sur le mois de juillet, d'estimer les passages par le sentier des muletiers à partir des données de comptages routiers.

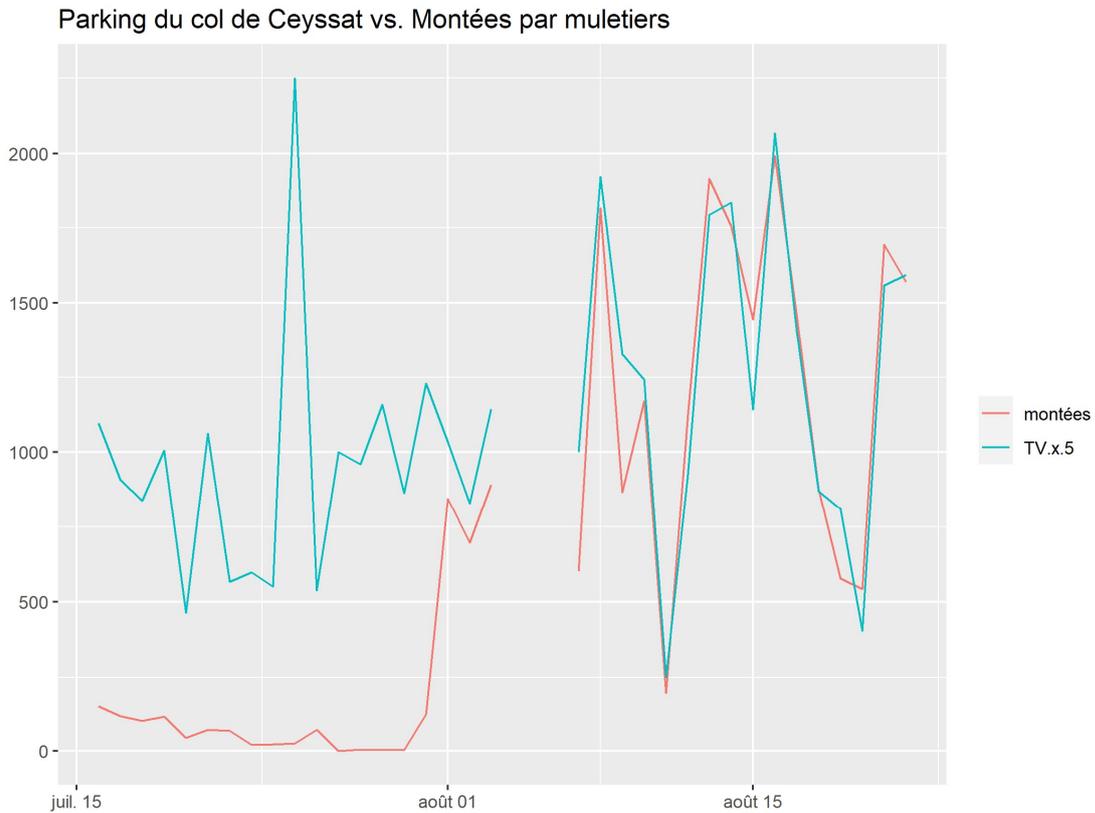


Figure 51 - Nombre de montées par l'éco-compteur des muletiers (rouge) et nombre de véhicules ayant stationné sur le parking du col de Ceysat (x5) entre le 17 juillet et le 22 août

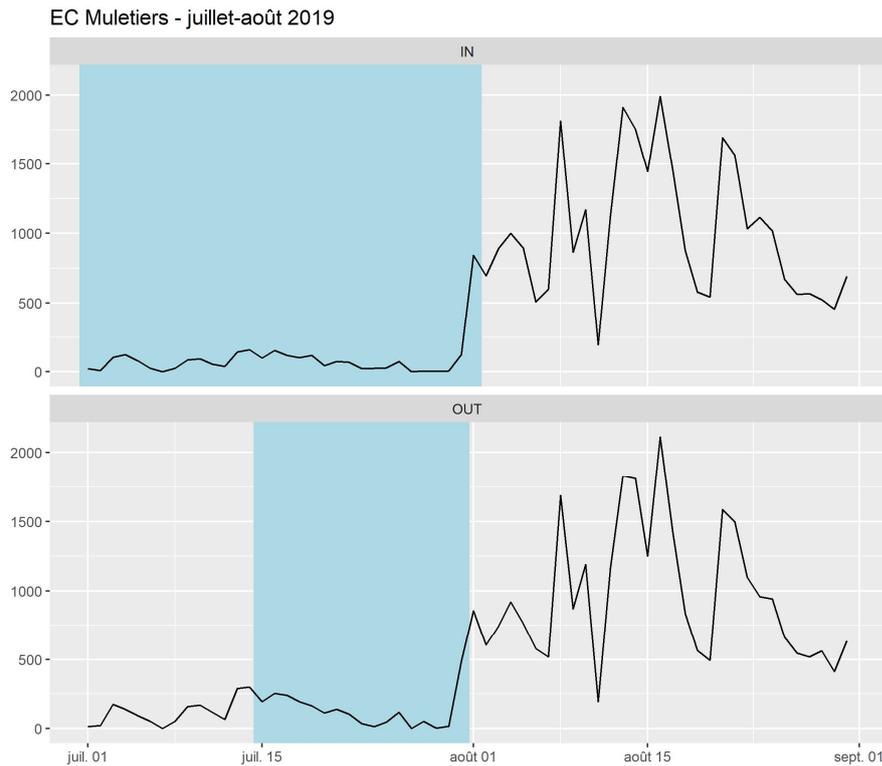


Figure 52 - Séquences anormalement basses en juillet sur les données de l'éco-compteur du sentier des muletiers

### 3.1.2. COMPTAGES VIDEO

Un test de comptage vidéo a été réalisé le 14 août 2019 au niveau de l'entrée du parking de Lemptégy, par un prestataire, pour la Direction des Routes, qui nous a transmis ces données. Elles consistent en une mesure des différents flux par analyse d'images, sous la forme de matrices, avec une matrice par tranche d'un quart d'heure et par type de véhicule.

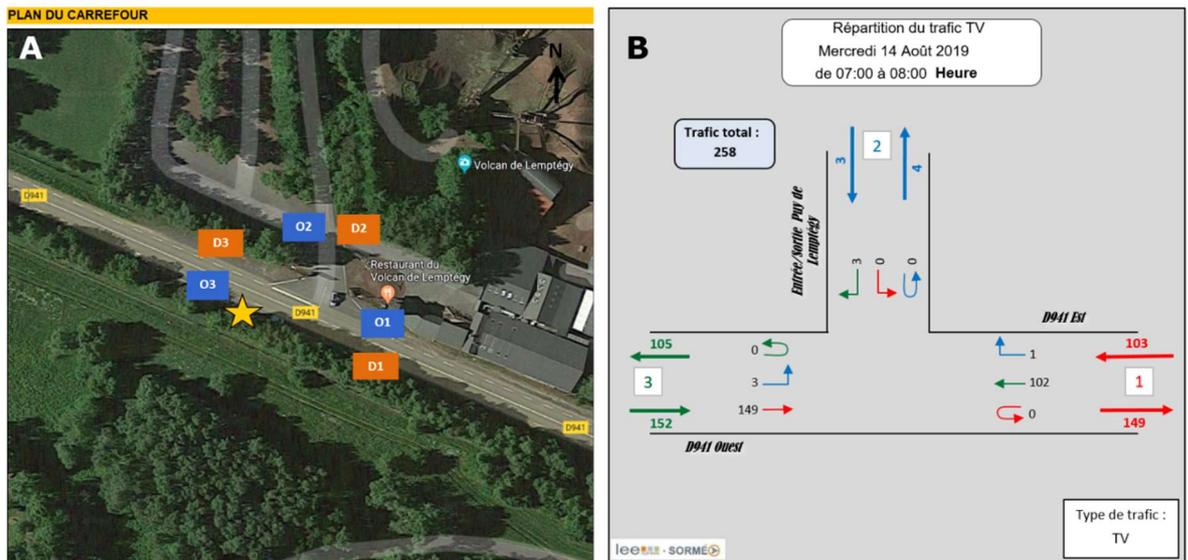


Figure 53 - Photo de l'entrée du parking du Volcan de Lemptégy avec les différentes voies (A) et schéma des flux mesurés par vidéo (B) – source : Lee Sormea (groupe Vinci)

A partir de ces données, il est possible de reconstituer les entrées et sorties du parking de Lemptégy, et ainsi de connaître le nombre de véhicules qui sont entrés dans le parking (Figure 53). A partir de ce chiffre, et du nombre d'entrées sur Lemptégy, on peut, par exemple, en déduire un nombre moyen de personnes par véhicule, en faisant l'hypothèse que tous les véhicules entrants sont des visiteurs du site (Figure 54). Cette estimation sous-estime probablement le chiffre réel, car il peut y avoir des personnes venues uniquement se renseigner, et d'autres qui sont déposées puis récupérées, sans compter les employés du site.

### Vidéo - Lemptegy - 14 août 2019 (VL)

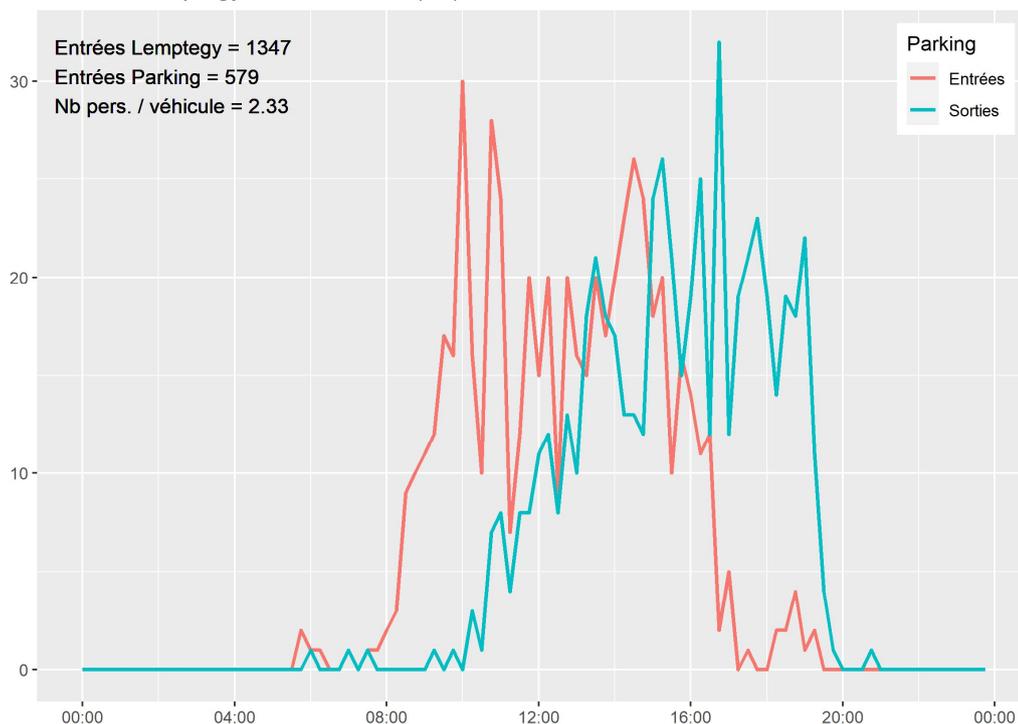


Figure 54 - Séries des entrées et sorties du parking du Volcan de Lemptegy le 14/08/2019

## 3.2. FLUX VISION

### 3.2.1. POURQUOI TESTER LE SERVICE FLUX VISION D'ORANGE ?

Début 2019, nous avons réfléchi à remplacer les comptages ponctuels sur les parkings, initialement prévus dans le projet, relativement lourds, coûteux, probablement difficiles à exploiter et sans doute moyennement fiables, par l'expérimentation d'un autre mode de comptage. Le recensement des techniques utilisées aujourd'hui dans le suivi de la fréquentation, notamment dans le secteur touristique, nous a conduit à envisager la solution Flux Vision d'Orange (Berlingué & Contet, 2017), développée à partir de 2012/2013 (Brémond, 2014), et déjà utilisée par plus de la moitié des ADT en 2015 (Villain, 2016).

Nous nous sommes donc renseignés sur ce service développé par Orange et, en effet, Flux Vision présentait a priori plusieurs avantages pour répondre à la problématique du suivi de la fréquentation sur la zone CPFL :

- Il fournit directement une estimation globale de l'ensemble des visiteurs présents sur la zone, sans risque de double compte, contrairement à une méthode basée sur l'agrégation de passages et d'entrées sur les différents sites ;
- L'estimation porte sur l'ensemble des visiteurs de la zone (avec une durée de présence cumulée sur la journée supérieure à 2h), et prend donc en compte des présences plus diffuses, ou en dehors des sites suivis (sites payants ou équipés d'éco-compteurs) ;
- Les données sont au grain journalier (et même infra-journalier, par tranche de 2h) ;

- Les données sont ventilées par type de visiteurs, selon une segmentation développée en concertation avec le réseau des ADTs, permettant de distinguer : touristes, excursionnistes, résidents... ;
- La méthode, développée depuis plus de 5 ans, se présente aujourd'hui comme stabilisée et reproductible dans le temps, permettant donc un suivi pluriannuel ;
- Les données fournies, obtenues à partir d'un échantillonnage massif sur l'ensemble des cartes SIM Orange, sont accompagnées d'informations riches sur les visiteurs : données socio-démographiques (âge, sexe, catégories socio-professionnelles), origines (départements, pays) ; et sur leur séjour (lieu de nuitée, durée de séjour, récurrence) ;
- Il est possible de suivre plusieurs zones d'étude en même temps, ainsi que des sous-zones, voire d'étudier les mobilités entre ces zones ;
- Les coûts semblent abordables, en comparaison des coûts d'une enquête, et sont largement dégressifs pour plusieurs années de suivi et/ou plusieurs zones ;
- Enfin, l'utilisation par les services de tourisme de très nombreuses collectivités locales constitue une certaine caution.

Au préalable, nous avons eu plusieurs échanges, notamment avec Jean-François Jobert (01/02/2019) et avec Guilhem Poncy (07/02/2019) au CDT Auvergne-Rhône-Alpes, qui a ont fait part de la mauvaise expérience de la région AURA avec Flux Vision, lorsque la région avait fait l'acquisition de deux années de données Flux Vision, entre 2014 et 2016. Ils nous ont signalé notamment les anomalies détectées à l'époque, entre les grandeurs connues et celles estimées par Flux Vision, notamment des écarts très importants concernant les visiteurs étrangers sur la région Auvergne-Rhône-Alpes.

Comme cette expérience était relativement ancienne, à une époque où Flux Vision était probablement encore en développement, nous avons voulu avoir d'autres retours portant sur des expériences plus récentes. Nous avons eu un long échange avec Gilles Chappaz (21/03/2019), précédemment en charge des données Flux Vision pour le Parc National des Calanques, qui nous a fait part de l'intérêt de ces données dans les suivis réalisés, plus récemment, dans le Parc National des Calanques. Il nous a toutefois signalé plusieurs limites rencontrées, notamment liées au débords (ex. détection des portables situés sur des navires de croisière passant au large). Il nous a indiqué que le service avait évolué depuis 2016, et avait probablement corrigé certains défauts, notamment concernant l'estimation des visiteurs étrangers.

Pariant sur l'amélioration du service depuis 2016, nous avons décidé de tester cette approche, principalement en raison des angles morts de l'approche par site (nombreux secteurs non couverts par des systèmes de comptage : sentiers non-équipés par des éco-compteurs, plage d'Aydat...), de l'importance de prendre en compte la fraction diffuse de la fréquentation, difficilement quantifiable autrement, et enfin en raison de la difficulté de corriger le problème de double comptes dans l'agrégation des données par site. Flux Visio pourrait ainsi produire une estimation plus exhaustive de la fréquentation, mais aussi sans risque de compter plusieurs fois les mêmes visiteurs.

Cette expérimentation permettrait à la fois d'apporter un élément de comparaison à la solution agrégée, obtenue à partir des sources actuellement disponibles, mais également de tester la fiabilité de ce service, qui pourrait être une approche alternative de suivi de la fréquentation, dans la mesure où ce service est aujourd'hui très largement utilisé, présenté comme une méthode reproductible, et potentiellement plus exhaustif (tourisme diffus, lieux non équipés), non redondant (pas de double comptes des personnes visitant plusieurs lieux ou

passant par plusieurs éco-compteurs dans la même journée), et beaucoup plus riche en terme de contenu (données socio-démographiques, lieux de nuitées...).

Enfin, plusieurs échanges avec Etienne Rames (CDT Occitanie) en décembre 2020 et janvier 2021, ainsi qu'avec Mireille Carniel-Fabre (ADT Hérault) le 8 mars 2021, ont apporté des éclairages supplémentaires sur la solution Flux Vision, de la part de services utilisant Flux Vision depuis des années.

### 3.2.2. PRESENTATION DE FLUX VISION

#### PRINCIPE

Le principe de Flux Vision repose sur la détection des cartes SIM Orange (portables allumés) dans une zone à partir des signaux détectés par les antennes proches, ainsi que des téléphones étrangers en roaming sur le réseau Orange (Figure 55). Les données sont échantillonnées en continu (et anonymisées), puis alimentent en temps réel des compteurs, à partir desquels seront calculés les indicateurs fournis par Orange, après plusieurs redressements successifs (Figure 56). Un premier redressement permet de passer d'un nombre de mobiles détectés à un nombre réel de mobiles présents, notamment à partir des parts de marché d'Orange, pour représenter l'ensemble du marché de la téléphonie mobile. Ensuite, à partir des données démographiques (INSEE), un second redressement vise à être représentatif de la population standard. Enfin un ultime redressement permet de prendre en compte des contextes et spécificités locales.

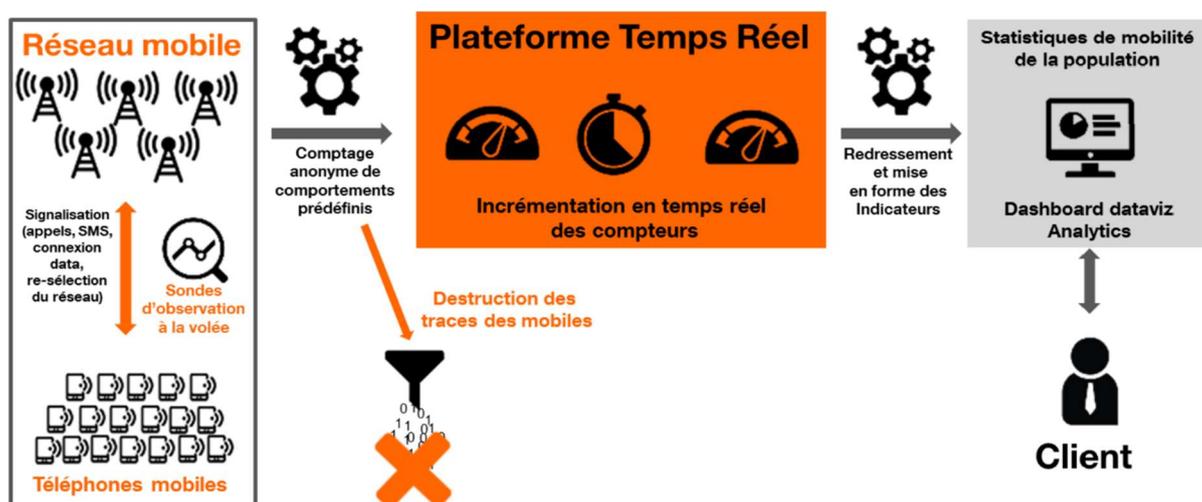


Figure 55 - Schéma simplifié du fonctionnement de Flux Vision (source : Orange, documentation Flux Vision)

## Détails sur les étapes du redressement

Afin de garantir la représentativité de la donnée, plusieurs étapes de redressement sont mises en place :

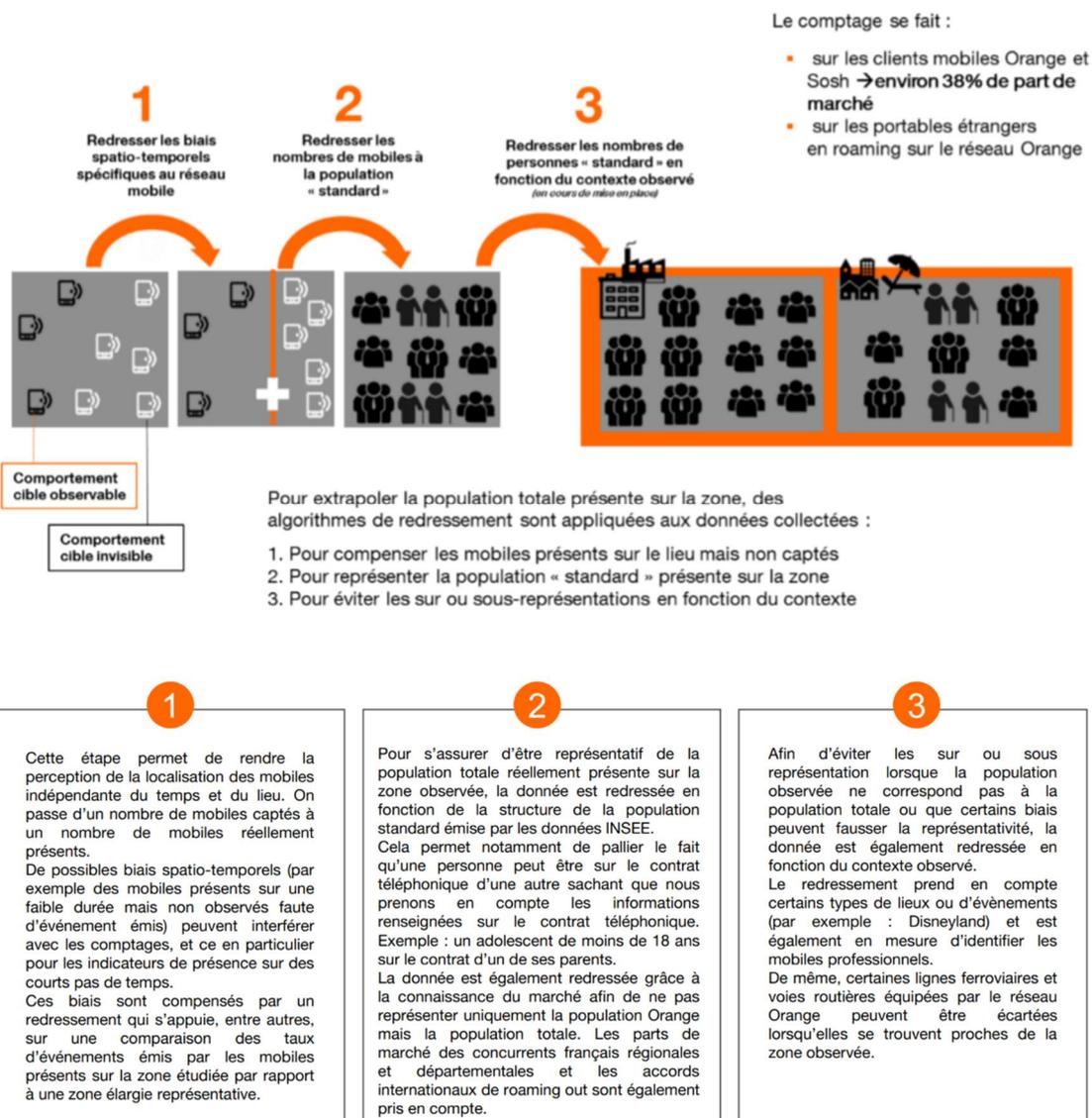


Figure 56 – Étapes du redressement des données dans Flux Vision (source : Orange, documentation Flux Vision)

## SEGMENTATION

Une segmentation des visiteurs a été mise au point en concertation avec le réseau des ADT. Il a abouti à un découpage en 9 classes, fonction du lieu de présence majoritaire, du lieu de nuitée, du temps de présence et de la récurrence sur la zone d'étude (Figure 57). Ainsi on distingue d'abord les personnes étant restées moins de 2 heures en cumulé sur la zone dans la même journée, qui sont définies comme « en transit ». Parmi les personnes restées plus de 2 heures en cumulé sur la zone dans la même journée, les « dormants » désignent les personnes ayant passé plus de 2h dans la zone entre 0h et 6h (et en présence majoritaire sur la zone). Parmi les « dormants » (Tableau 7), on définit les « résidents » comme des personnes, dont le lieu de présence majoritaire entre 0h et 6h se situe dans la zone, et qui ont passé plus de 22 nuits dans

la zone pendant les 8 dernières semaines. Les autres « dormants » sont des « habituellement présents » s'ils ont été vus dans la zone au moins 4 fois sur des semaines différentes au cours des 8 dernières semaines ; dans les autres cas, il s'agit de « touristes » au sens de Flux Vision. Les « non-dormants » (Tableau 8) désignent les « excursionnistes » et les « excursionnistes récurrents », ces derniers ayant été vus au moins 5 fois dans la zone au cours des 15 derniers jours.

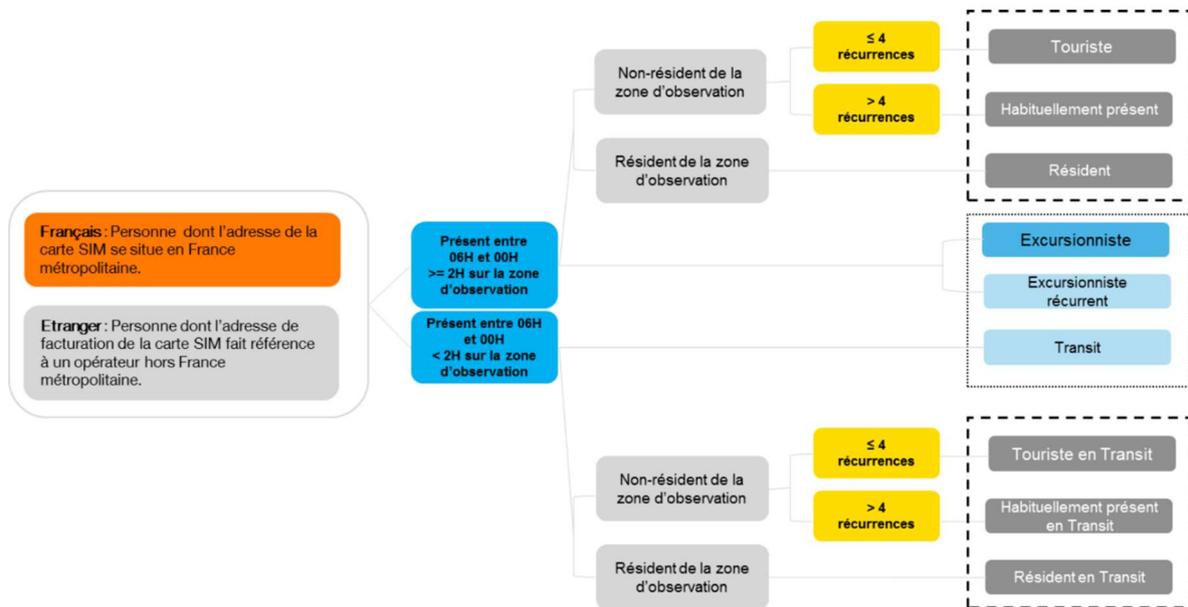


Figure 57 - Segmentation des visiteurs dans Flux Vision Tourisme

Segmentation des dormants sur la zone d'observation :

- **Résident** : Personne dont la zone de présence majoritaire (lieu où la personne a passé le plus de temps entre 00H et 06H sur toute la France métropolitaine) est dans la zone d'observation et qui a réalisé **au moins 22 nuits (pas forcément consécutives)** dans la zone de présence majoritaire (au sein de la zone d'observation)
- **Habituellement présent** : Personne qui n'est pas résidente de la zone d'observation mais qui a été vue de manière **récurrente** sur cette zone : **plus de 4 fois sur des semaines différentes au cours des 8 dernières semaines.**
- **Touriste** : Personne qui n'est ni **résidente**, ni **habituellement présente**. C'est-à-dire
  - La zone d'observation n'est pas sa zone de résidence (zone de présence majoritaire sur laquelle la personne doit passer au moins 22 nuits)
  - La personne a été vue de manière non récurrente sur cette zone : au maximum sur 4 semaines (au moins une fois par semaine) sur les 8 dernières semaines.

Tableau 7 - Segmentation des "dormants" (source :Orange, documentation Flux Vision)

Segmentation des non dormants de la zone d'observation :

- **Excursionniste** : Personne présente plus de 2H sur la zone d'observation et ne réalisant pas de nuitées la veille et le soir du jour étudié et **présente moins de 5 fois en journée sur les 15 derniers jours.**
- **Excursionniste récurrent** : Personne présente plus de 2H sur la zone d'observation et ne réalisant pas de nuitées la veille et le soir du jour étudié et présente **plus de 5 fois en journée sur les 15 derniers jours.**

**Transit** : Personne présente moins de 2H sur la zone d'observation et ne réalisant pas de nuitées la veille et le soir du jour étudié.

Tableau 8 - Segmentation des "non dormants" (source :Orange, documentation Flux Vision)

---

## DEFINITION DE LA ZONE

La définition de la zone d'étude consiste à rechercher la meilleure concordance entre la zone d'étude et la zone, obtenue par différentes simulations réalisées par Orange (4 types), représentant des probabilités élevées de détection. Comme la concordance n'est jamais parfaite, il s'agit de trouver le meilleur compromis possible.

Vous trouverez ci-dessous des exemples de couverture avec en noir la zone souhaitée et, en couleur, la couverture analysable par Flux Vision :

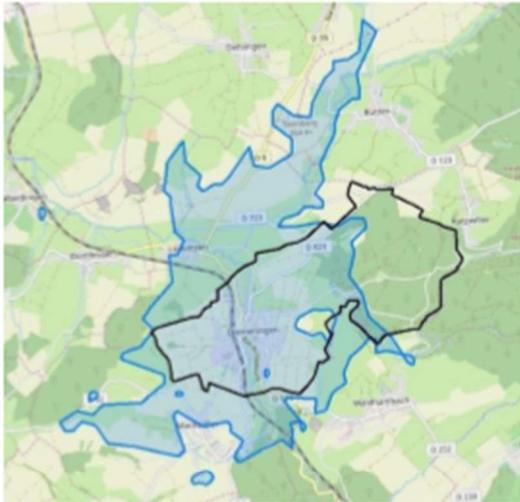


Figure 1 : zone souhaitée non conciliable avec le réseau d'antenne

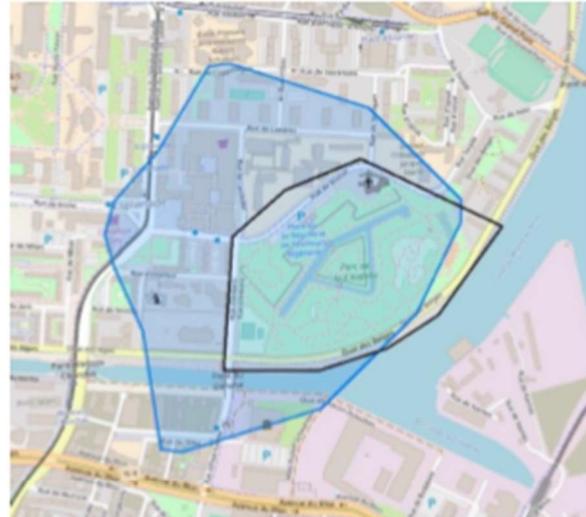


Figure 2 : zone souhaitée trop réduite

Selon la densité géographique des antennes et la taille des zones du territoire à étudier, la qualité de cette localisation peut varier. Dans certains cas défavorables, certaines zones devront être agrandies car elles sont difficilement dissociables par le réseau mobile.

Figure 58 - Problèmes de définition de la zone d'étude (source : Orange, documentation Flux Vision)

---

## INTERPOLATIONS

Il peut arriver que les données Flux Vision soient impactées par une indisponibilité partielle ou complète des sondes du réseau Orange. La livraison comprend un fichier qui renseigne la disponibilité, pour chaque jour et chaque nuit, en utilisant le code suivant : 1 pour « incidence partielle ou totale », 0 pour « Aucune problématique de capture ». Dès qu'une indisponibilité est constatée, des valeurs interpolées sont proposées sur les zones impactées. Ces interpolations sont construites à partir d'un modèle statistique prenant en compte les données disponibles et des données externes (variables météorologiques et calendaires).

---

## INTERFACE

Pour les usages standards de Flux Vision, une interface PowerBI est disponible, offrant un tableau de bord avec représentations graphiques des indicateurs, et des outils de reporting.

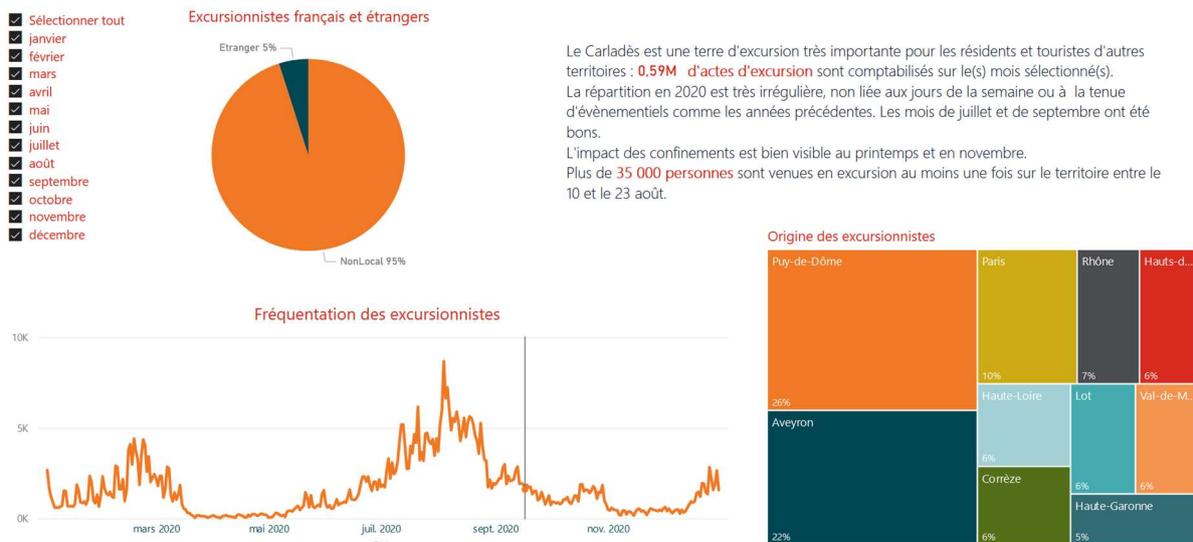


Figure 59 - Exemple de visualisation des données via l'interface powerBI de Flux Vision (département du Cantal)

## COUT

Le montant de l'achat réalisé en 2019 s'est élevé à 8 190 euros HT, pour 3 mois de données Flux Vision Tourisme sur 4 zones différentes, plus un mois de l'offre Flux Vision Géomarketing sur une zone (sous-zone Vulcania). En revanche, le prix annoncé pour une année complète sur une seule zone, était d'environ 10 000 euros, tarif dégressif pour un suivi plus long ou d'avantage de zones.

D'après les échanges que l'on a eus avec d'autres organismes ou collectivités, le tarif est négociable et très dégressif. Ainsi nous a-t-on parlé de montants de l'ordre de 25 000 euros pour 30 zones sur 2 ans.

### 3.2.3. ZONES D'ETUDE

Le choix des zones d'études est complexe. Il a nécessité de multiples allers et retours avec Orange, et plusieurs réunions, afin de comprendre le principe et arriver à trouver des compromis plus ou moins satisfaisant. En effet, à partir de nos requêtes, Orange fournit des cartes qui sont des simulations correspondant à des probabilités élevées de détection. Nous avons essayé différentes options, mais la marge de manœuvre est en réalité assez limitée, en particulier dans cette zone, contrainte par le relief et la couverture très limitée.

Fin juillet 2019, nous avons fait l'acquisition de 3 mois de données Flux Vision Tourisme, portant sur la période du 1<sup>er</sup> août au 31 octobre 2019, sur 4 zones :

- Zone A : une zone complète étendue, couvrant toute la zone du bien et la zone tampon, avec des débords principalement à l'ouest, et un peu au sud, de la zone (Figure 60) ;
- Zone B : une zone complète réduite, centrée sur la Chaîne des Puys jusqu'à Aydat, mais ne couvrant pas la partie la plus au Nord, à proximité de Volvic (Figure 61) ;
- Zone C : une sous-zone centrée sur Lemptégy et Vulcania, présentant juste quelques débords vers Bromont-Lamothe (Figure 62) ;
- Zone D : une sous-zone centrée sur la zone Puy de Dôme – Puy de Pariou (Figure 63).

Ce qui a guidé notre choix des deux premières zones, compte tenu des contraintes lourdes posées par la définition de ces zones et des contraintes budgétaires liées au projet, fut de disposer en quelques sortes d'une estimation haute avec la zone complète étendue (zone A), et d'une estimation basse de la fréquentation avec la zone complète réduite (zone B). D'autre part, n'ayant pas d'expérience de Flux Vision, nous n'avons aucune assurance par rapport à l'importance réelle des débords sur les données. En effet, Orange définit ces cartes comme des probabilités fortes de détection, mais nous n'avons aucune idée et aucune possibilité de connaître la définition précise de ces probabilités. Lors de nos échanges avec Flux-Vision, nous n'avons pas pu avoir d'informations claires sur d'éventuelles fluctuations de ces frontières, ni sur l'importance et l'étendue des détections au-delà des zones cartographiées. C'est pourquoi, nous avons choisi d'intégrer une zone potentiellement « exhaustive », et une zone plus réduite, « de sécurité », limitées aux principaux sites, afin d'avoir le minimum de problèmes liés aux débords, si ces derniers devaient s'avérer difficiles à traiter.

Pour les deux sous-zones, l'idée était de disposer de zones de test, pour juger de la fiabilité de la mesure Flux Vision. En effet, d'après les cartes des zones C et D, ces zones ne couvrent pas, ou presque pas, de zones urbanisées, et ne devraient donc pas intégrer de présences autres que les visiteurs de tourisme et de loisirs, parmi les présences cumulées de plus de deux heures. En outre, compte tenu des sources de données potentiellement disponibles, ces sous-zones permettaient de réaliser des tests ciblés, en particulier s'agissant de la zone Lemptégy-Vulcania, où les données pourraient simplement être confrontées au cumul des billetteries des deux sites, dans la mesure où a priori seuls les visiteurs ayant pris un billet d'entrée seraient susceptibles de rester suffisamment longtemps pour être comptabilisés par Orange. En ayant accès aux données journalières de billetterie, même sur le seul mois d'août 2019, et idéalement accompagnées de l'origine des visiteurs, cette sous-zone semblait particulièrement intéressante pour réaliser un test de la fiabilité de Flux Vision.

En complément, nous avons souhaité testé un autre service développé par Orange, **Flux Vision Géomarketing**, qui repose sur le même principe que FVT, mais plus ciblé sur la fréquentation de lieux commerciaux. Il se distingue de FVT par l'absence de segmentation touristique, mais aussi par une segmentation par durée de présence, avec deux seuils : 30 minutes, et 3 heures de présence cumulée. Cette commande porte uniquement sur la zone Lemptégy-Vulcania, sur le mois d'août 2019, et visait également à compléter le test de fiabilité de la méthode Flux Vision, notamment au cas où les présences cumulées à Lemptégy pouvait être inférieure à 2 heures.

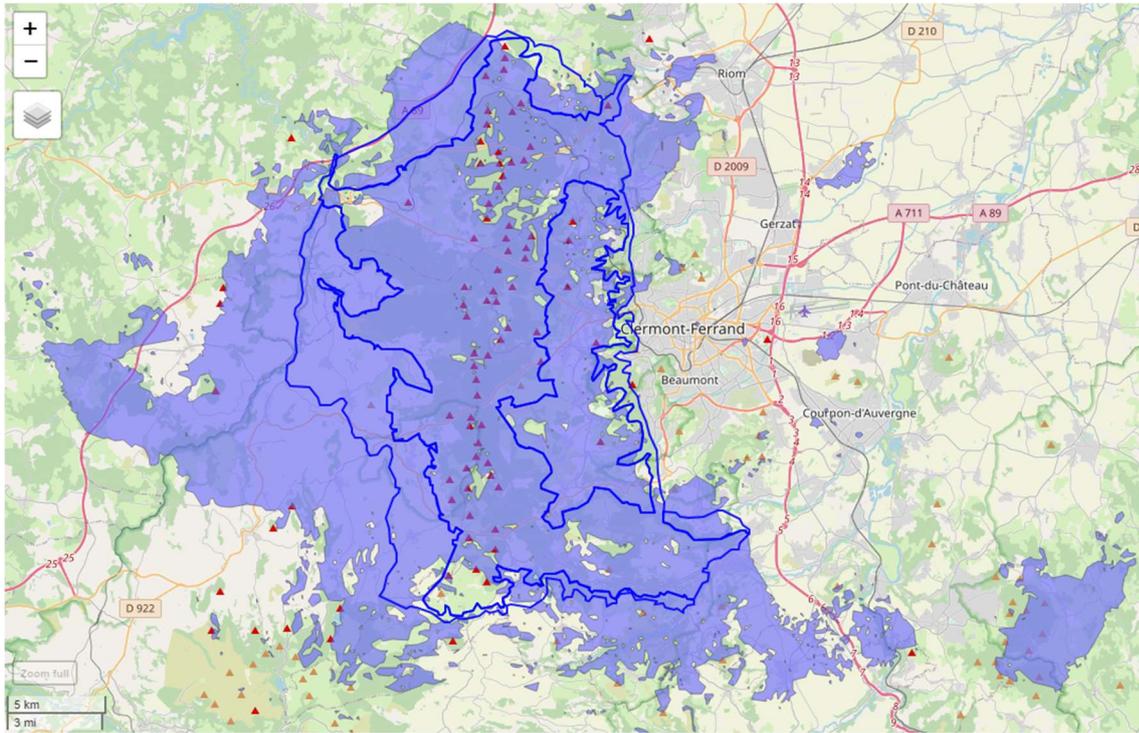


Figure 60 - Zone A (zone complète étendue)

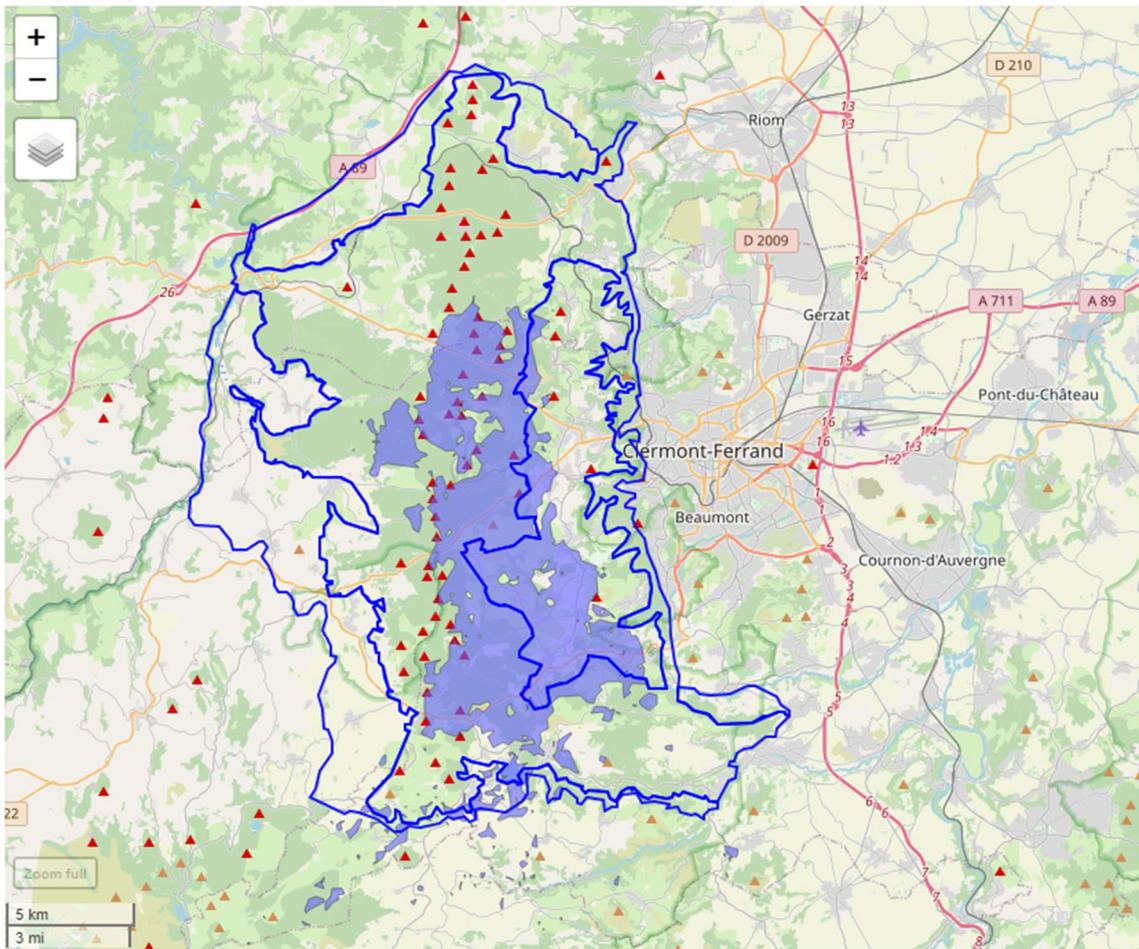


Figure 61 - Zone B (zone complète réduite)

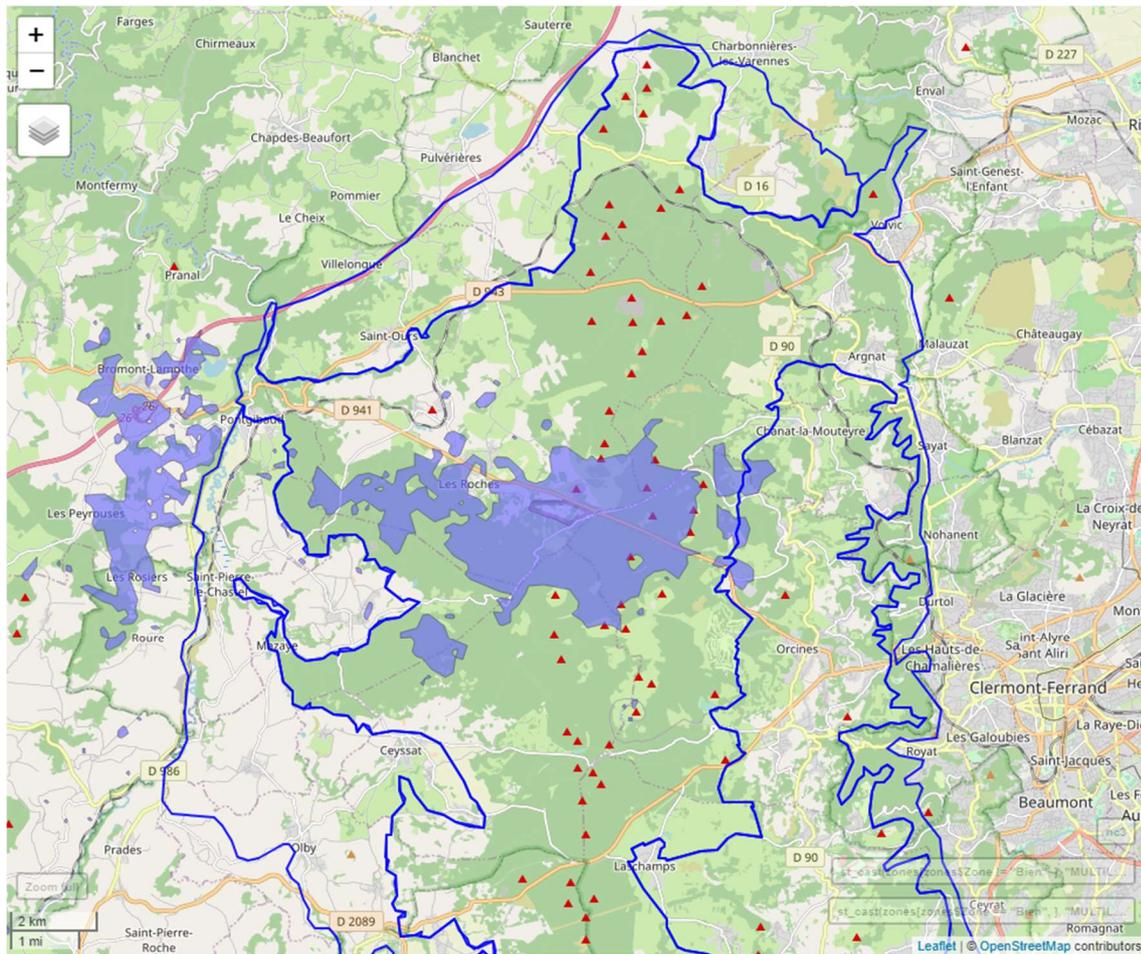


Figure 62 - Zone C (sous-zone centrée sur Lemptéguy et Vulcania)

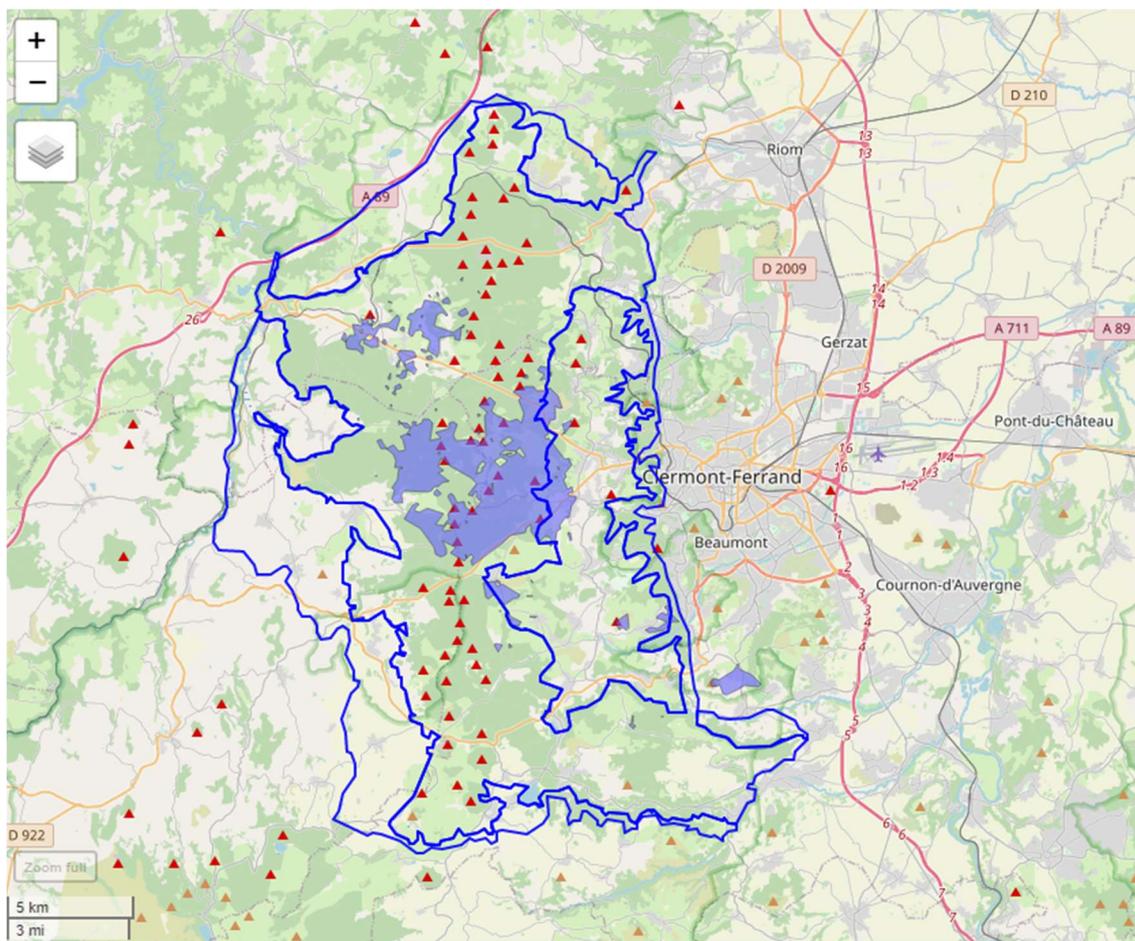


Figure 63 - Zone D (sous-zone centrée sur Pariou – Puy de Dôme)

### 3.2.4. DEUX OFFRES FLUX VISION

#### OFFRE FLUX VISION TOURISME (FVT)

L'offre Flux Vision Tourisme (FVT) est le service spécifiquement dédié au tourisme. L'essentiel des données commandées en 2019 sont des données Flux Vision Tourisme.

Sur ces données, à partir des échantillons redressés, sont calculés toute une série d'indicateurs :

- Arrivées / Départs (dormants) ;
- **Présence diurne (> 2h) ;**
- Fréquentation (tranche 2h) ;
- Lieu de présence majoritaire (dormants) ;
- **Lieu de nuitée ;**
- Nombre de nuitées (dormants) ;

- Récurrence ;
- Durée de séjour (dormants).

Ainsi, pour chaque zone, Flux Vision Tourisme fournit 60 tableaux de données (Tableau 9). Ces tableaux fournissent des informations journalières sur les nombres de personnes présentes plus de 2h en cumulé sur la zone, portant sur :

- Les arrivées correspondant aux dormants sur la zone non vus les deux nuitées précédentes (Arrivees) ;
- Les départs, correspondant aux dormants sur la zone non vus les deux nuitées suivantes (Departs) ;
- Les présences en journées de plus de 2h en cumulé sur la tranche 06h-00h (Diurne) ;
- Les durées de présence en journée (Duree\_Presence) ;
- Les présences par tranche de 2h (Frequentation\_2H) ;
- Les lieux de présence majoritaire des dormants sur la zone (Mobilite\_Diurne\_Nuitees) ;
- Les lieux de nuitée des personnes présentes en journée (Mobilite\_Nuitees) ;
- Les nombres de nuitées (Nuitees) ;
- La récurrence ou nombre de fois où la personne présente en journée a été vue sur la zone au cours des 15 derniers jours (Recurrence) ;
- La durée de séjour des dormants, écart entre départ et arrivée (Duree\_Sejour) ;
- La présence d'une indisponibilité partielle ou totale des sondes du réseau Orange conduisant à une interpolation avec les données disponibles et des données externes (Dispo).
- Les données de ces indicateurs sont également ventilées selon d'autres segmentations : classes d'âge (Age), sexe (Civilite), catégorie socio-professionnelle (Geolife), pays d'origine (Pays), département d'origine (Regions).

La plupart de ces indicateurs sont ventilés par :

- Nationalité (étranger/français) ;
- Type de visiteur (segmentation FVT) ;
- Lieu de facturation (Intra/Extra Puy de Dôme).

Et, certains, par :

- Caractéristiques socio-démographiques : sexe, âge, typologie Géolife ;
- Origine (Pays, département).

Les indicateurs de mobilité sont disponibles pour les nuitées sur zone la veille (Nuitees\_Veille) ou le soir (Nuitees\_Soir) du jour de présence. Enfin sont également disponibles les cumuls par mois des présences en journée et des nuitées sur zone.

 Arrivees	 Diurne_Mois_Geolife	 Mobilite_Nuitees_Veille
 Arrivees_Age	 Diurne_Mois_Pays	 Mobilite_Nuitees_Veille_Pays
 Arrivees_Civilite	 Diurne_Mois_Regions	 Mobilite_Nuitees_Veille_Regions
 Arrivees_Geolife	 Diurne_Pays	 Nuitees
 Arrivees_Pays	 Diurne_Regions	 Nuitees_Age
 Arrivees_Regions	 Duree_Presence	 Nuitees_Civilite
 Departs	 Duree_Presence_Pays	 Nuitees_Geolife
 Departs_Age	 Duree_Presence_Regions	 Nuitees_Mois
 Departs_Civilite	 Frequentation_2H	 Nuitees_Mois_Age
 Departs_Geolife	 Frequentation_2H_Pays	 Nuitees_Mois_Civilite
 Departs_Pays	 Frequentation_2H_Regions	 Nuitees_Mois_Geolife
 Departs_Regions	 Mobilite_Diurne_Nuitees_Soir	 Nuitees_Mois_Pays
 Dispo	 Mobilite_Diurne_Nuitees_Soir_Pays	 Nuitees_Mois_Regions
 Diurne	 Mobilite_Diurne_Nuitees_Soir_Regions	 Nuitees_Pays
 Diurne_Age	 Mobilite_Diurne_Nuitees_Veille	 Nuitees_Regions
 Diurne_Civilite	 Mobilite_Diurne_Nuitees_Veille_Pays	 Recurrence
 Diurne_Geolife	 Mobilite_Diurne_Nuitees_Veille_Regions	 Recurrence_Pays
 Diurne_Mois	 Mobilite_Nuitees_Soir	 Recurrence_Regions
 Diurne_Mois_Age	 Mobilite_Nuitees_Soir_Pays	 Sejours_Duree
 Diurne_Mois_Civilite	 Mobilite_Nuitees_Soir_Regions	 Sejours_Duree_Pays

Tableau 9 - Liste des fichiers de données disponibles dans l'offre Flux Vision Tourisme

## OFFRE GEOMARKETING (GMK)

Le service Flux Vision Géomarketing se distingue de l'offre Flux Vision Tourisme d'une part par l'absence de segmentation touristique (Tableau 10), mais aussi par la distinction de plusieurs seuils de durée de présence cumulée (30 minutes et 3 heures, contre seulement un seuil de 2 heures pour FVT). Cette segmentation par la durée de présence permet d'identifier :

- Les présences furtives, inférieure à 30 minutes (considérées comme du transit) ;
- Les présences significatives entre 30 minutes et 3 heures (considérées comme des vrais visiteurs) ;
- Et les activités importantes sur la zone, supérieure à 3 heures (considérées comme des résidents ou des personnes travaillant sur la zone, dans les études géomarketing).

Dans notre étude, l'idée était de tester la fiabilité de cet autre service, avec également l'idée que les durées de présence dans les sites payants, en particulier au Volcan de Lemptégý, pouvaient éventuellement être inférieures au seuil de 2 heures de Flux Vision Tourisme. Ainsi, dans l'optique d'un test de la fiabilité des estimations de Flux Vision, réalisé sur la sous-zone Vulcania-Lemptégý, il était nécessaire de s'assurer que toutes les visites (payantes) seraient

bien prises en compte. En faisant l’hypothèse que les personnes venant se renseigner ou transitant par la zone restaient moins d’une demi-heure, et que les personnes visitant le Volcan de Lemptéguy, plus petit que Vulcania, pouvaient dans certains cas rester moins de 2 heures sur le site, on pourrait peut-être avoir une estimation meilleure avec le seuil de 30 minutes de GMK qu’avec FVT.

 entrees	 presencejournee
 presence30min	 presencejournee-activite-bassindevie
 presence30min-activite-bassindevie	 presencejournee-activite-communes
 presence30min-activite-communes	 presencejournee-activite-iris
 presence30min-activite-iris	 presencejournee-activitenuitee-bassindevie
 presence30min-destination-bassindevie	 presencejournee-activitenuitee-communes
 presence30min-destination-communes	 presencejournee-activitenuitee-iris
 presence30min-destination-iris	 presencejournee-cp
 presence30min-nuitee-bassindevie	 presencejournee-nationalite
 presence30min-nuitee-communes	 presencejournee-nuitee-bassindevie
 presence30min-nuitee-iris	 presencejournee-nuitee-communes
 presence30min-origine-bassindevie	 presencejournee-nuitee-iris
 presence30min-origine-communes	 sorties
 presence30min-origine-iris	

Tableau 10 - Liste des fichiers de données disponibles dans l’offre Flux Vision Géomarketing

### 3.2.5. DIAGNOSTIC DES DONNEES

#### ZONES D’ETUDE

Comme on l’a vu, le choix a été fait, par sécurité, de retenir deux zones : l’une couvrant la totalité de la zone CPFL (bien et zone tampon), avec des débords principalement vers l’ouest et le moins possible vers Clermont-Ferrand (Figure 64-A) ; l’autre, beaucoup plus réduite, couvrant les zones les plus touristiques, c’est-à-dire l’essentiel des puys jusqu’à Aydat (Figure 64-B). L’idée était d’avoir, au mieux, deux estimations, une estimation basse et une estimation haute de la fréquentation ; et, au pire, de disposer, avec la zone réduite, d’une solution de secours, au cas où les débords soient beaucoup plus impactant que prévu par les cartes.

Deux sous-zones ont également été commandées, afin de réaliser des tests plus précis de fiabilité du service Flux Vision, en confrontant ces données aux données disponibles (éco-compteurs, billetteries) : l’une centrée sur Vulcania et Lemptéguy (Figure 64-C), l’autre sur la zone Puy de Dôme – Puy de Pariou (Figure 64-D).

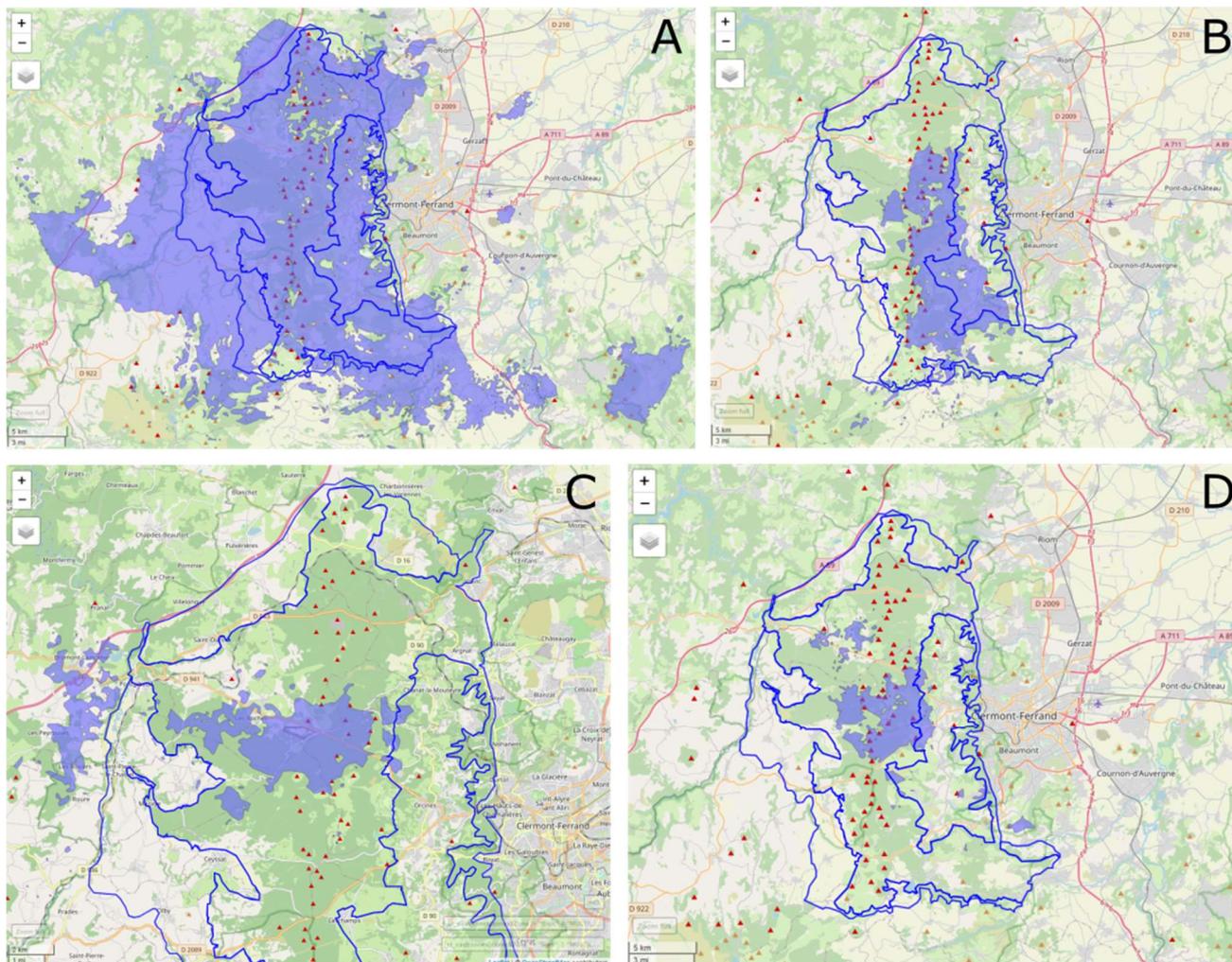


Figure 64 - Zones d'étude Flux Vision (A = Zone complète étendue, B = Zone complète réduite, C = Sous-zone centrée sur Vulcania, D = Sous-zone centrée sur le Puy de Dôme)

Une première vérification de la validité des mesures Flux Vision a consisté à confronter le nombre de résidents estimés par Flux Vision sur les zones d'étude à la population recensée par l'INSEE. Pour ce faire, les données carroyées de l'INSEE (données filosofi 2015, carreau de 200 m de côté et 4 ha de superficie) (INSEE, 2019) ont été croisées avec les polygones des cartes de probabilité de présence Flux Vision, afin de calculer le cumul de population résidente attendue sur l'emprise de chacune des zones. On obtient ainsi une population théorique que l'on compare au nombre de résidents Flux Vision, estimés par le nombre de résidents journalier moyen, sur une plage où ce nombre est a priori stationnaire, soit hors vacances scolaires ; ici entre mi-septembre et mi-octobre.

A partir des résultats (Tableau 11), on peut déjà constater que :

- Sur la zone complète étendue : les écarts sont importants en valeur absolue en raison probablement de nombreux débords en ville et dans les zones résidentielles, rendant les corrections sans doute difficiles ;
- Sur la zone complète réduite : les écarts sont certes importants en relatifs, mais beaucoup plus limités en valeur absolue, permettant peut-être une correction des problèmes liés aux débords ;

- Sur la sous-zone Puy de Dôme : les écarts sont énormes, et les valeurs Flux Vision totalement aberrantes, ne permettant pas d'exploiter ces données ;
- Sur la sous-zone Vulcania : il y a très peu de résidents et les valeurs Flux Vision restent limitées, avec seulement quelques centaines de résidents détectés, ce qui doit correspondre à des débords limités sur des petites communes proches, ne risquant sans doute pas de générer beaucoup de « faux visiteurs », faisant donc de cette zone une zone de test très intéressante (avec en outre 2 offres disponibles : Flux Vision Tourisme et GéoMarketing).

Zone	INSEE	Flux Vision
Zone complète étendue	55 354	76 764
Zone complète réduite	4 341	8 453
Sous-zone Puy de Dôme	1 862	67 015
Sous-zone Vulcania	107	601

Tableau 11 - Résidents Flux Vision et population INSEE sur les zones Flux Vision (sources : filosofi 2015, Flux Vision)

## AUTRES ANOMALIES

Un premier travail a consisté à explorer les nombreux tableaux fournis par Flux Vision, à bien comprendre l'ensemble des concepts, parfois complexes, voire peu intuitifs, tels que la définition des types de visiteurs en fonction de jours de présence sur les semaines précédentes, la segmentation glissante, la récurrence, les entrées et sorties, les nuitées veille et soir, le transit, la présence nocturne, la zone de facturation, les interpolations, les groupements de communes spécifiques à Flux Vision, le filtrage RGPD... La bonne compréhension de tous ces points a nécessité de bien étudier les documentations Flux Vision, d'analyser en profondeur les nombreux tableaux de données, ainsi que des allers-retours fréquents avec Orange, par mail et en visio.

Au fil des analyses exploratoires, plusieurs anomalies sont apparues, pour lesquelles il n'y avait pas toujours d'explication évidente, même pour les experts d'Orange.

En premier lieu, des débords plus importants que ceux attendus d'après les cartes ont pu être mis en évidence à partir des « dormants » ayant des nuitées dans des communes hors zone.

Ensuite, à partir des tableaux correspondant aux lieux de nuitée, des communes avec des nombres de nuitées d'excursionnistes habitant le Puy de Dôme plus élevés qu'attendu ont été mis en évidence, écarts dont il est en revanche difficile d'établir la ou les cause(s) : visites dans les débords (amis/famille, boutiques...), déplacements aller/retour à travers la zone, comptabilisés comme des temps de présence cumulés supérieur à 2h...

Enfin, on peut également citer un dernier point, encore moins explicable : la part très importante de résidents en transit, c'est-à-dire de résidents passant moins de 2h en cumulé sur la zone d'étude (donc la zone de résidence) sur la plage 06h-00h, alors que cette part devrait a priori être très faible (en effet, entre 6h et minuit, la plupart des résidents ont passé au moins 2h à leur domicile). D'après Orange, cela pourrait être dû à des événements émis par le téléphone vers des antennes 4G plus lointaines, conduisant à des localisations hors-zone.

De même, en raison des spécificités de la zone (faible réseau d’antennes, relief), on constate des phénomènes d’entrées-sorties de « dormants », qui conduisent à des indicateurs inexploitable, notamment pour les durées de séjour, qui sont anormalement courtes.

Ces derniers résultats ne sont pas particulièrement rassurants quant à la fiabilité des autres mesures, pour lesquelles il n’y a pas d’anomalie évidente.

Enfin, d’autres difficultés, non explicitement renseignées dans la documentation Flux Vision, apparaissent au cours des analyses, tels que la non concordance des effectifs cumulés sur les différents tableaux. Par exemple, le nombre de visiteurs avec lieux de nuitée renseignés est nettement inférieur au nombre de visiteurs total, ce qui nécessite parfois de redresser les données d’effectifs avec nuitées (par une règle de trois).

### 3.2.6. CORRECTIONS

#### PROBLEMES LIES AUX DEBORDS

Un problème majeur provient des débords, c’est-à-dire des détections en dehors de la zone d’étude. De plus, le travail exploratoire a mis en évidence des débords au-delà de ce qui était attendu d’après les cartes Flux Vision. Selon l’importance de ces débords et le type de zones concernées, le biais peut être plus ou moins grave. En effet, si les débords couvrent une zone résidentielle, une zone urbaine avec différents services, on peut s’attendre à des flux de visiteurs non concernés par la zone d’étude. De même, si des attractions ou des hébergements touristiques sont présents dans les débords, des flux de touristes ne se rendant pas forcément sur la zone d’étude seront pris en compte. Pour appréhender ces problèmes, un indicateur tel que le ratio, pour chaque groupement de communes, des nombres journaliers de nuitées sur la population totale permet de faire un premier diagnostic (Tableau 12) ; si ce nombre est trop élevé, on peut penser que des visiteurs de ces communes ne se rendent pas réellement dans la zone d’étude (mais plutôt dans les débords).

En outre, d’autres caractéristiques de la zone CDPFL complexifient les problèmes liés aux débords, notamment :

- La proximité de la métropole de Clermont-Ferrand et de sa périphérie ;
- Le transit important à travers la zone.

**Tableau 12 - Regroupements de communes Flux Vision sur le Puy de Dôme avec distance à la zone CDPFL, population, nombre d’hébergements, et ratio nuitées sur population**

CP	Communes	Distance	Population	Hébergements	Nuitées	Ratio
<b>NEBOUZAT</b>	NEBOUZAT	0	161	834	40,5	4,85
<b>LA BOURBOULE</b>	LA BOURBOULE	11 619	1716	1 782	59,8	3,36
<b>SAINT-DIERY</b>	COURGOUL / CRESTE / SAINT-DIERY / SAINT-PIERRE-COLAMINE / VERRIERES	8 137	188	866	28,5	3,29
<b>MAZAYE</b>	CEYSSAT / MAZAYE / SAINT-PIERRE-LE-CHASTEL	0	15	1 849	48,4	2,62

<b>MUROL</b>	MUROL / LE VERNET-SAINTE-MARGUERITE	231	1 618	882	22,2	2,52
<b>CHAMBON-SUR-LAC</b>	CHAMBON-SUR-LAC	6012	1240	410	9,8	2,38
<b>MONTEL-DE-GELAT</b>	COMBRAILLES / CONDAT-EN-COMBRAILLE / LANDOGNE / MONTEL-DE-GELAT / TRALAIGUES / VILLOSANGES	12 410	175	1 744	40,6	2,33
<b>PERPEZAT</b>	BRIFFONS / HEUME-L'EGLISE / LAQUEUILLE / PERPEZAT / TORTEBESSE	4730	73	1236	28,5	2,31
<b>PARENT</b>	PARENT	9 231	25	841	17,8	2,11
<b>CHAMALIERES</b>	CHAMALIERES	0	84	17 282	333,7	1,93
<b>VEYRE-MONTON</b>	VEYRE-MONTON	1 468	0	3 489	67,2	1,93
<b>MONT-DORE</b>	MONT-DORE / ORCIVAL / ROCHEFORT-MONTAGNE / VERNINES	0	1724	2 841	54,5	1,92
<b>GIAT</b>	FERNOEL / GIAT / HERMENT / VERNEUGHEOL	18 089	74	1 461	23,7	1,62
<b>GELLES</b>	GELLES / OLBY	0	108	1 743	28,2	1,62
<b>SAINTE-SATURNIN</b>	COURNOLS / OLLOIX / SAINT-SATURNIN	0	125	1 722	26,5	1,54
<b>DURTOL</b>	DURTOL / NOHANENT	0	18	4 235	63,3	1,49
<b>AUBIERE</b>	AUBIERE / CHANONAT / ROMAGNAT	0	356	19 490	265,8	1,36
<b>SAINTE-SAUVES-D'AUVERGNE</b>	MURAT-LE-QUAIRE / SAINT-JULIEN-PUY-LAVEZE / SAINT-SAUVES-D'AUVERGNE / TAUVES	8 780	882	2 749	36,2	1,32
<b>LES MARTRES-DE-VEYRE</b>	AUTHEZAT / CORENT / LES MARTRES-DE-VEYRE / MIREFLEURS	3 917	108	7 797	100,0	1,28
<b>BEAUMONT</b>	BEAUMONT / CEYRAT / ROYAT	0	934	22 146	277,4	1,25
<b>BROMONT-LAMOTHE</b>	BROMONT-LAMOTHE	0	145	974	11,4	1,17
<b>LE CENDRE</b>	LE CENDRE / DALLET / MEZEL / PERIGNAT-SUR-ALLIER / LA ROCHE-NOIRE / SAINT-BONNET-LES-ALLIER / SAINT-GEORGES-SUR-ALLIER	5 199	0	9 130	97,3	1,07
<b>MESSEIX</b>	AVEZE / MESSEIX / SAINT-SULPICE / SINGLES	17 248	288	1 482	15,7	1,06
<b>LA TOUR-D'AUVERGNE</b>	LA TOUR-D'AUVERGNE	13 245	405	644	6,5	1,01
<b>BOURG-LASTIC</b>	BOURG-LASTIC / LASTIC / SAINT-GERMAIN-PRES-HERMENT	16 788	7	1 070	10,8	1,01
<b>PLAUZAT</b>	CHAMPEIX / CLEMENSAT / GRANDEYROLLES / LUDESSE / MONTAIGUT-LE-BLANC / PLAUZAT / SAINT-FLORET / SAURIER	2 006	294	5 006	39,7	0,79
<b>PROMPSAT</b>	PROMPSAT / TEILHEDE	5 101	0	860	6,8	0,79
<b>LE BREUIL-SUR-COUZE</b>	LE BREUIL-SUR-COUZE / LE BROU / NONETTE / ORSONNETTE / LES PRADEAUX / SAINT-MARTIN-DES-PLAINS / SAINT-REMY-DE-CHARGNAT	18 031	329	2 777	20,7	0,75
<b>LA ROCHE-BLANCHE</b>	LE CREST / LA ROCHE-BLANCHE / SAINT-AMANT-TALLENDE / SAINT-SANDOUX / TALLENDE	0	0	8 855	65,1	0,74
<b>SAINTE-GEORGES-DE-MONS</b>	SAINTE-GEORGES-DE-MONS	3 422	129	1 976	13,6	0,69

<b>MOZAC</b>	MARSAT / MOZAC	1 843	0	5 221	34,6	0,66
<b>COURNON-D'AUVERGNE</b>	COURNON-D'AUVERGNE / ORCET / PERIGNAT-LES-SARLIEVE	2 846	658	25 438	163,5	0,64
<b>CLERMONT-FERRAND</b>	CLERMONT-FERRAND	340	2 519	142 686	900,4	0,63

## ALGORITHME DE CORRECTION

Pour pallier les problèmes de faux visiteurs liés aux débords, nous avons développé plusieurs algorithmes de correction des données FVT, qui ont été testés successivement, le principe sous-jacent restant toutefois le même.

En premier lieu, il s'agit, à partir des données de lieux de nuitées, d'identifier des groupements de communes pour lesquels les nombres de nuitées moyens sont a priori trop importants, c'est-à-dire des groupements de communes ayant potentiellement des interactions avec les débords de la zone Flux Vision (déplacements dans les communes des débords, ou peut-être communes en limite des débords avec détection partielle en journée).

Pour ce faire, un indicateur (R) mesurant le rapport entre le nombre moyen de nuitées et le potentiel de nuitée a priori calculé à partir de la population et du nombre d'hébergements des communes du groupement (Figure 65).

Dans un second temps, on ne considère plus que les groupements de communes dans un rayon de 50 km autour de la zone CPFL, et on distingue deux groupes parmi les groupements de communes, sur la base de l'indicateur R, calculé précédemment. On identifie un seuil  $R_0$ , à partir de la distribution de l'indicateur R, correspondant à une rupture nette dans la distribution (Figure 66), permettant de séparer les groupements de communes en deux sous-ensembles :

- Les GC avec des nombres de nuitées moyens anormalement élevés, soit  $R > R_0$  (GCA) ;
- Les GC avec des nombres de nuitées moyens a priori normaux, soit  $R < R_0$  (GCN).

Dans un troisième temps, on modélise le nombre de nuitées attendu en fonction de plusieurs variables :

- La distance minimale du GC à la zone CPFL ;
- La population du GC ;
- Le nombre d'hébergements du GC.

La population et le nombre d'hébergements sont calculés par agrégation des valeurs des communes du groupement (données INSEE).

Le nombre d'hébergements (heb) est calculé ainsi :

$$heb = coll * 25 + hotel + camp * 2.5$$

où *hotel*, *camp* et *coll* représentent respectivement le nombre de lits d'hôtel, le nombre d'emplacements de camping, et le nombre d'autres hébergements collectifs de tourisme.

Le modèle de régression est ajusté sur les GC avec des nombres de nuitées moyens normaux, sans Clermont-Ferrand, pour éviter un trop fort effet de levier et parce que les communes concernées sont beaucoup plus petites.

A partir de ce modèle, on prédit la valeur attendue du nombre de nuitées dans les autres GC, ceux avec des nombres de nuitées anormalement élevés. La somme des écarts entre les nombres observés et attendus de nuitées correspond à la fraction à corriger. Enfin, comme le nombre de nuitées total est inférieur au nombre total de visiteurs, on redresse cette valeur par le ratio des effectifs totaux des deux tables (présences en journée et nuitées).

$$\left( \sum_{GCA} (Observé_i - Prédit_i) \right) \times \frac{Nb\ total\ de\ visiteurs}{Nb\ total\ de\ visiteurs\ avec\ lieu\ de\ nuitée}$$

Plusieurs versions de cet algorithme ont été testées successivement afin d'affiner la correction, ainsi que plusieurs versions de l'indicateur, plusieurs seuils et plusieurs spécifications du modèle de régression.

Dans une première version de l'algorithme, les effectifs de touristes et d'excursionnistes étaient corrigés en même temps, à partir des nombres de nuitées moyens, avec un indicateur hybride, prenant en compte la population et l'hébergement touristique.

Dans une seconde version, les corrections étaient cette fois réalisées séparément pour les touristes et les excursionnistes, toujours sur les nombres de nuitées moyens.

Dans une troisième version, les mêmes corrections sont réalisées, séparément, mais cette fois sur les nombres de nuitées journaliers, après avoir identifié, sur les nombres moyens de nuitées, les GC avec des nombres de nuitées anormaux.

Enfin, dans une dernière version, celle utilisée ici, le même algorithme que le précédent est appliqué, mais sans correction des effectifs de touristes. En effet, en sélectionnant bien la fraction touriste, les différences de nombres de nuitées des touristes dans les différents GC n'étaient pas très marquées (pas de rupture nette de l'indicateur R) et, en outre, pouvaient peut-être s'expliquer par ailleurs. De plus, le risque de visites de plus de 2 heures dans les débords par des touristes est a priori faible ; ces visites longues dans les débords ne devraient donc concerner que des résidents des communes proches, donc des excursionnistes Flux Vision.

Après une correction, date par date, des données « excursionnistes », il est possible de reconstituer les séries complètes de données journalières pour les touristes, les excursionnistes et l'ensemble des visiteurs (touristes + excursionnistes), sur toute la période août-octobre 2019.

## Excursionnistes - Ratio Nuitées / Population

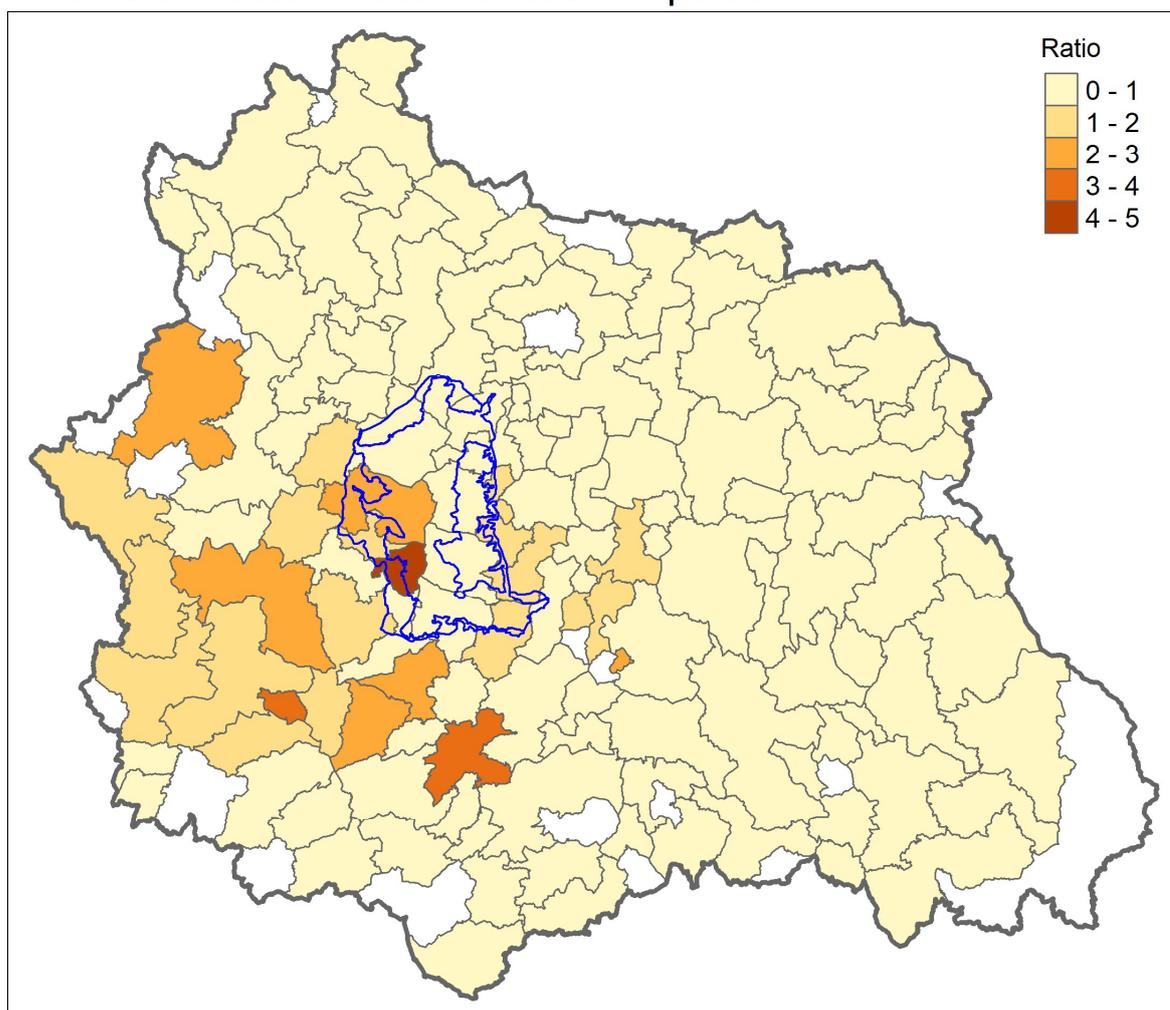


Figure 65 - Carte des groupements de communes selon les valeurs du ratio des nuitées « Excursionnistes » sur la population

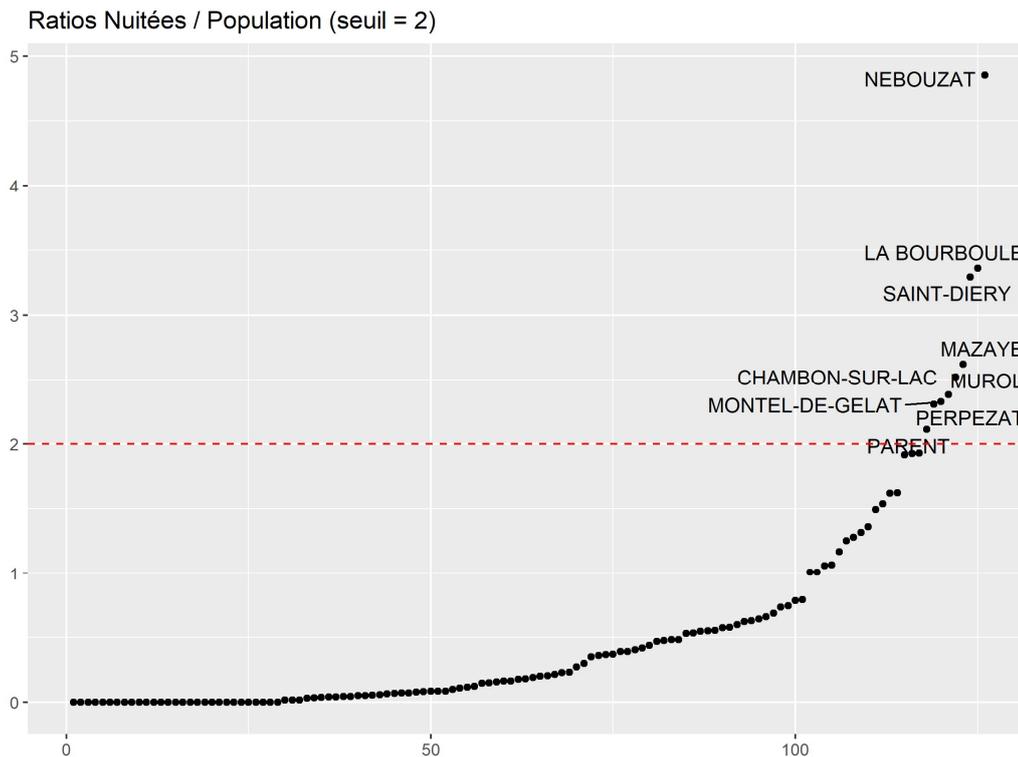


Figure 66 - Distribution des ratios Nuitées / Population, et choix d'un seuil

## 4. COMPARAISONS DES METHODES

### 4.1. COMPARAISON FLUX VISION / AGREGAT DES SITES DENOMBRABLES

#### 4.1.1. SOUS-ZONE LEMPTÉGY-VULCANIA

##### COMPARAISON DES SERIES

Pour tester la fiabilité des estimations de Flux Vision, l'achat de données sur une sous-zone centrée sur Vulcania et Lemptégy devait permettre de confronter ces données aux données de billetterie des deux parcs. Malheureusement, lorsque nous avons sollicité les deux entreprises, en novembre 2019, seul le Volcan de Lemptégy nous a fourni l'ensemble des données nécessaires pour ce test, à savoir les données journalières d'entrée pour le mois d'août 2019 (ventilées par origine). Vulcania a finalement accepté, en avril 2021, de nous transmettre un échantillon de 8 dates (sans les origines).

En outre, nous disposons pour cette zone, sur le mois d'août 2019, de deux offres Flux Vision : l'offre Flux Vision Tourisme (décrite précédemment), et l'offre Flux Vision Géomarketing (GMK), se distinguant, entre autres, par l'absence de segmentation touristique et la présence de deux seuils de présence cumulée (30 minutes et 3h). Le choix de tester cette offre supplémentaire était motivé par l'interrogation sur le fait que les visiteurs puissent ne pas

systématiquement rester plus de 2h (en particulier sur le site de Lemptégy, plus petit que Vulcania).

En première analyse, sur les données brutes (Figure 67), on peut déjà constater que le seuil de 30 minutes produit des estimations beaucoup trop élevées, incluant peut-être à des personnes venues uniquement se renseigner, voire du simple transit. Les autres courbes semblent à première vue assez proches et relativement bien corrélées.

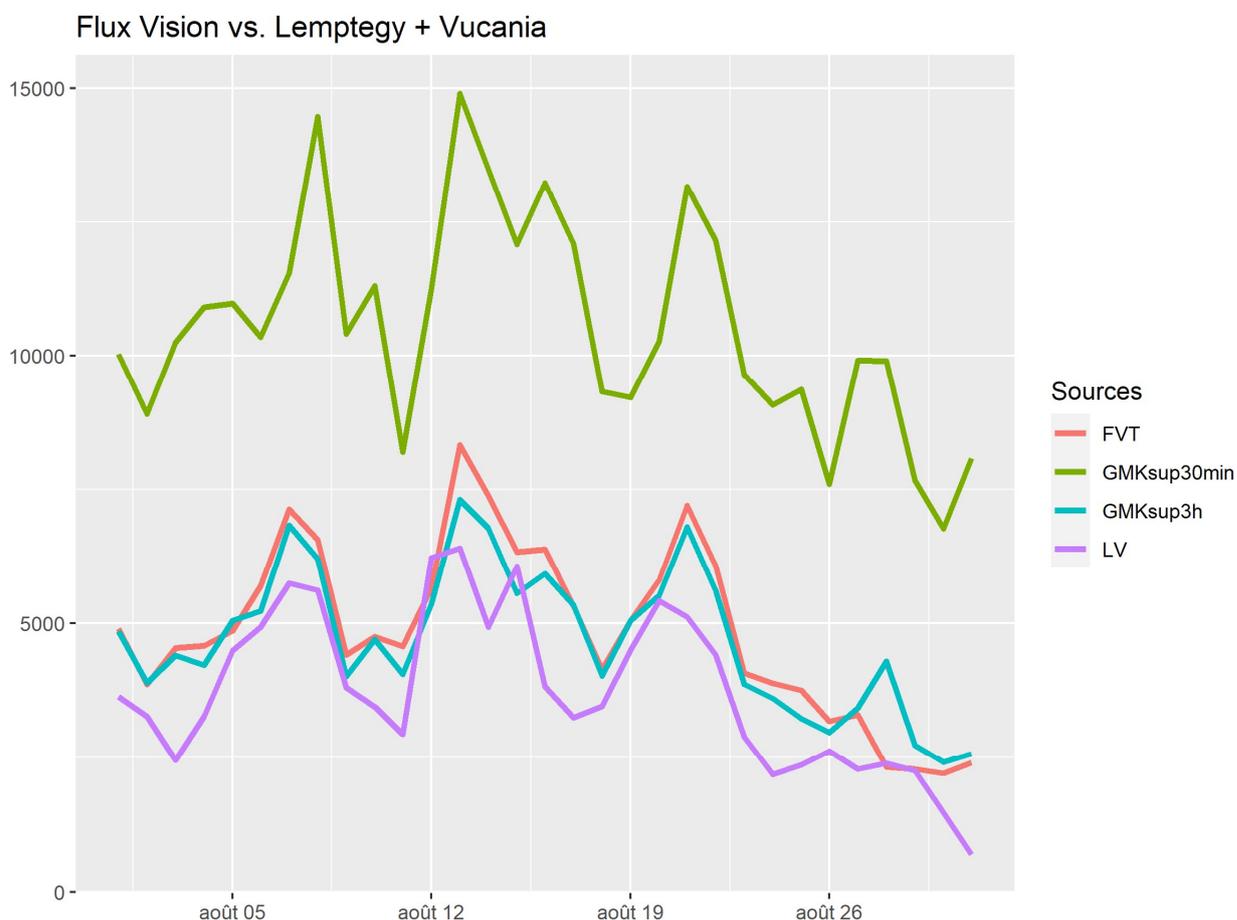


Figure 67 - Données journalières Flux Vision vs. Vulcania+Lemptégy (août 2019)

En ne regardant que les deux courbes FVT et GMK au seuil de 3h et les deux agrégats suivants :

- LV = Lemptégy + Vulcania (8 vraies valeurs, et valeur estimée à 3 fois Lemptégy pour les autres dates) ;
- LVP = LV + cumuls des passages extrapolés sur les 4 puys proches : Gouttes, Côme, Goules et Pariou.

on peut observer (Figure 68), que les 4 courbes sont très corrélées, qu'il y a un écart significatif entre la courbe LV et les 2 courbes Flux Vision (surtout FVT), que la courbe FVT est souvent très proches de la courbe LVP, notamment sur 4 des 8 dates, pour lesquelles on a les vraies entrées de Vulcania, moins sur les 2 valeurs interpolées (points rouges), sur deux dates autour du 10 août (mais sans vraies valeurs de Vulcania) et en fin de série, et enfin que la courbe GMK\_3h est plus proche de l'agrégat LV.

Il est difficile de savoir si l'offre Géomarketing produit des indicateurs plus centrés sur la zone cible ou si les pertes liées au seuil de 3h (ex. personnes restant moins de 3h à Lemptégy) sont compensées en partie par les personnes visitant les puys proches. D'autres éléments laissent penser que les débords excèdent la carte Flux Vision et couvrent tout ou partie des puys proches, et expliqueraient bien la proximité des courbes FVT (rouge) et LVP (violet). C'est pourquoi, il est dommage de ne pas disposer de l'ensemble des données Vulcania, qui aurait permis de trancher la question.

Flux Vision (FVT et GMK) vs. Agrégats Lemptégy + Vulcania + Puys

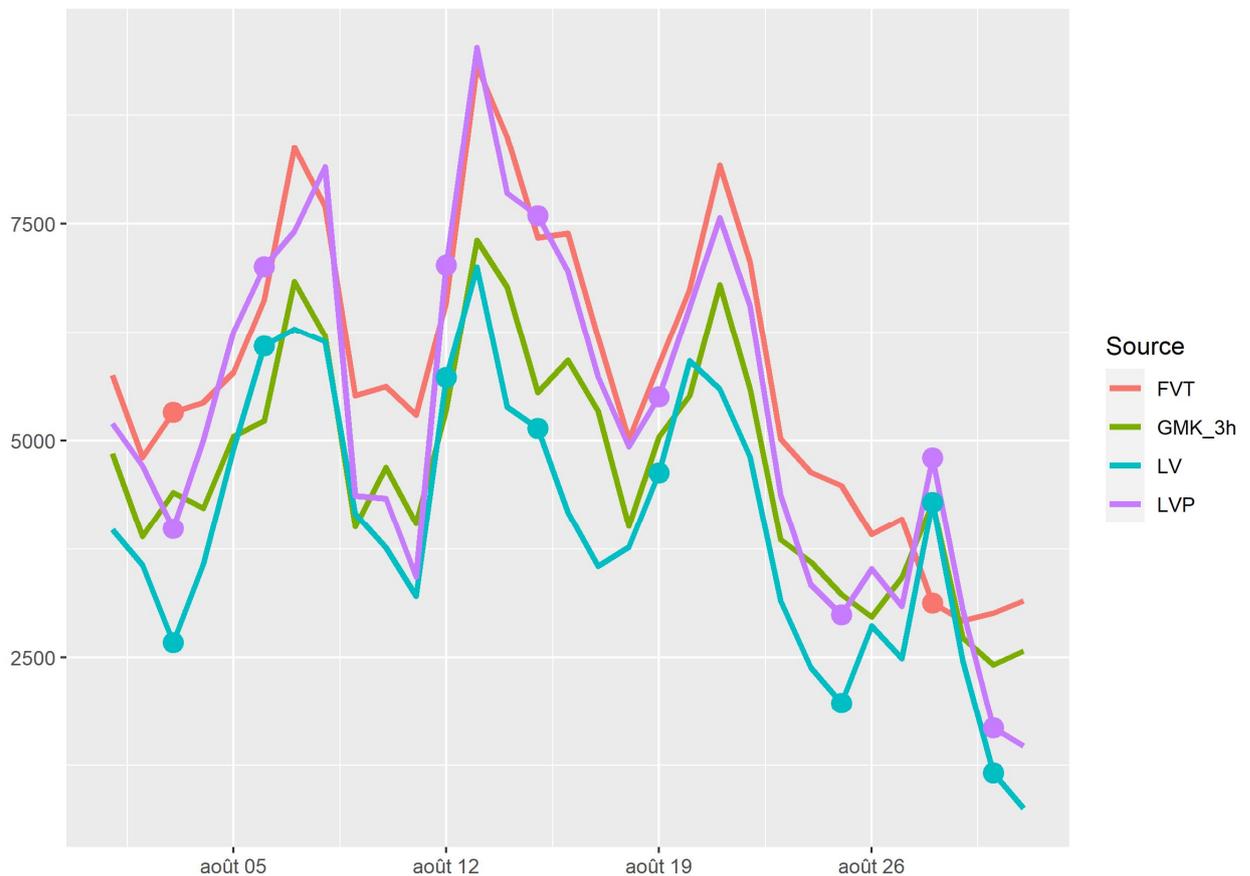


Figure 68 - Séries Flux Vision vs Agrégats billetteries + éco-compteurs (FVT = Flux Vision Tourisme, GMK\_3h = Flux Vision Géomarketing avec seuil de 3h, LV = cumul Lemptégy +Vulcania estimé, LVP = Lemptégy +Vulcania estimé + Puys de Gouttes, Goules, Côte et Pariou / points bleus et violet = données avec vraies valeurs pour Vulcania ; points rouges = données Flux Vision interpolées)

## ORIGINES DES VISITEURS

Si l'on regarde les proportions de visiteurs par origine (Tableau 13), après correction des données Flux Vision (effets de débord, cf. 3.2.6.), on constate que la proportion de visiteurs étrangers est un peu supérieure par rapport à Lemptégy, mais nettement inférieure à celle du Panoramique. Or, il est possible que pour Vulcania, cette proportion se situe entre celle de Lemptégy et celle du Panoramique. La proportion de visiteurs du Puy de Dôme est plus élevée, mais assez proche de celle du Panoramique pour FVT. Toutefois, il est difficile d'aller plus loin dans l'analyse, sans les données par origine de Vulcania.

Tableau 13 - Proportion de visiteurs par origine selon différentes sources

Source	Étrangers	D63
FVT	7.3 %	21.0 %
GMK_3h	6.7 %	24.9 %
Lemptégy	4.8 %	17.5 %
Panoramique	12.7 %	19.3 %

#### 4.2.2. ZONE COMPLETE REDUITE

Sur le mois d'août 2019, sur lequel on dispose des données les plus précises, la comparaison des données Flux Vision Tourisme corrigées (FVTc) à l'agrégat des sites dénombrables (SD), décrit dans la partie II, montre des séries très corrélées et très proches, à l'exception de quelques points, notamment sur la fin du mois (98). Ces résultats semblent plutôt valider les estimations Flux Vision, tout en retenant que les données de certains sites représentant une part importante de la fréquentation au mois d'août ne sont pas connues, et ici seulement grossièrement estimées (ex. Vulcania, Aydat). Là encore, la connaissance des données de Vulcania, et peut-être un raffinement de l'estimation sur Aydat, pourrait améliorer cette comparaison.

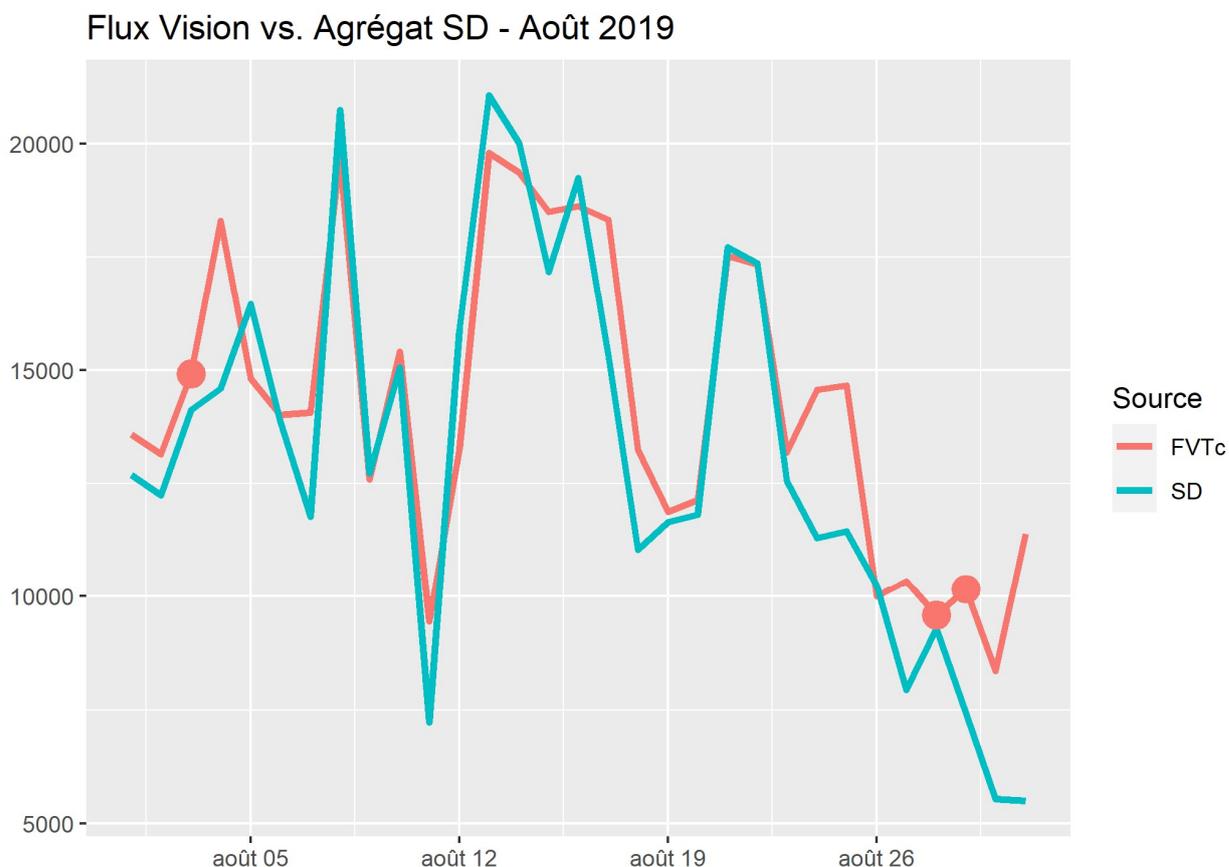


Figure 69 - Séries Flux Vision Tourisme corrigée et agrégat des sites dénombrables (sans Volvic) sur la ZCR au mois d'août 2019

Si l'on regarde cette fois l'ensemble de la série, sur les 3 mois allant d'août à septembre 2019 (Figure 70), les séries restent relativement synchrones, mais les écarts sont très importants pour les mois de septembre et octobre, bien que réduits les week-ends et durant les vacances de la Toussaint (fin octobre).

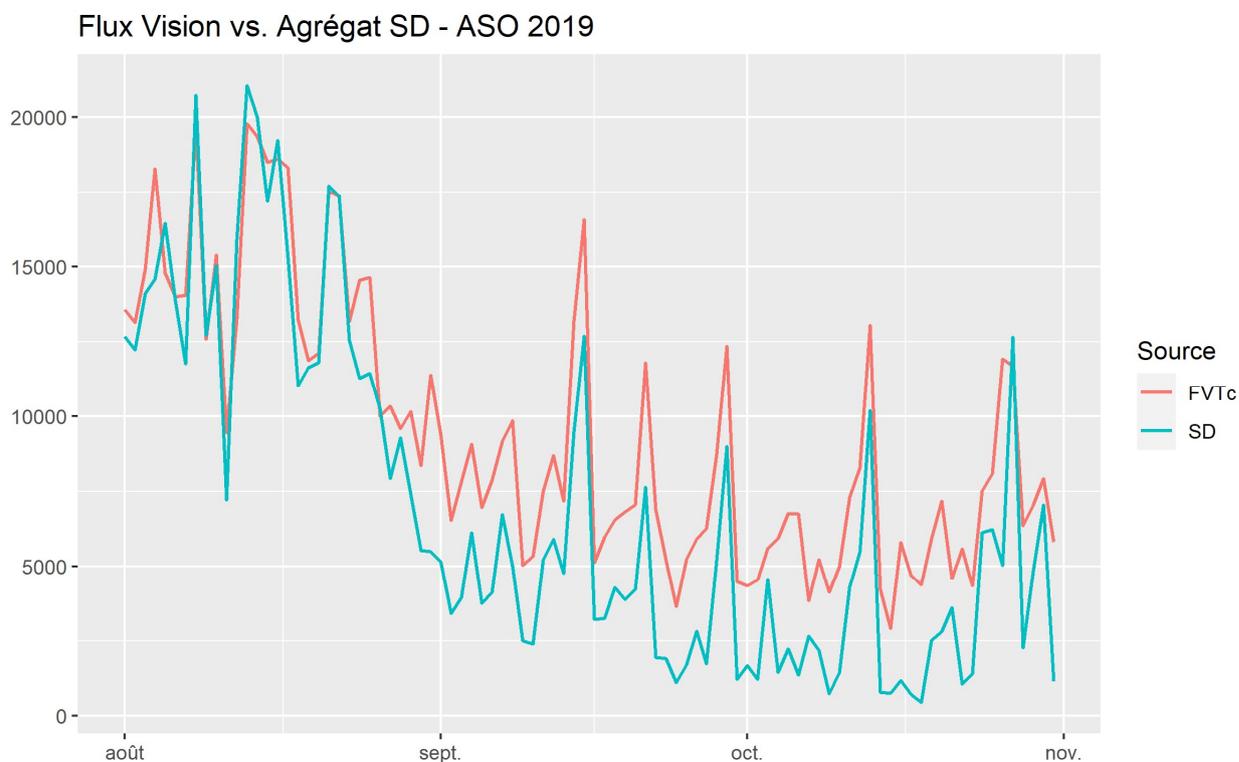


Figure 70 - Série Flux Vision Tourisme corrigée (FVTc) et agrégat des sites dénombrables (SD) sur la période août-octobre 2019

En ajoutant la seule composante « Touristes » de Flux Vision (Figure 71), on peut voir que celle-ci est très corrélée et très proche de la série SD sur les mois d'octobre et septembre, et inférieure les week-ends, lorsque les locaux sont davantage présents. Il y a donc une certaine cohérence de la composante touristique de Flux Vision avec les données des sites agrégées. Les écarts avec FVTc sont plus difficiles à expliquer. On peut s'attendre à ce que Flux Vision soit plus exhaustif, et intègre une part de la fréquentation qui n'est pas prise en compte dans l'agrégat SD : celle des sites et sentiers non-équipés, ainsi qu'une fréquentation diffuse (personnes circulant dans la zone CDPFL sans visiter de sites spécifiques, se promenant ou faisant du sport sur des sentiers annexes, ou venant pour se renseigner sur des sites, pour pique-niquer, aller au restaurant...). Une part notable de ces écarts peut correspondre à des pratiques plus diffuses (hors sites touristiques) d'excursionnistes locaux, plus importantes hors périodes de vacances. Enfin, il faut noter que les séries, notamment celles des sites payants, sont encore plus incertaines pour les mois de septembre et octobre.

## Flux Vision vs. Agrégat SD - ASO 2019

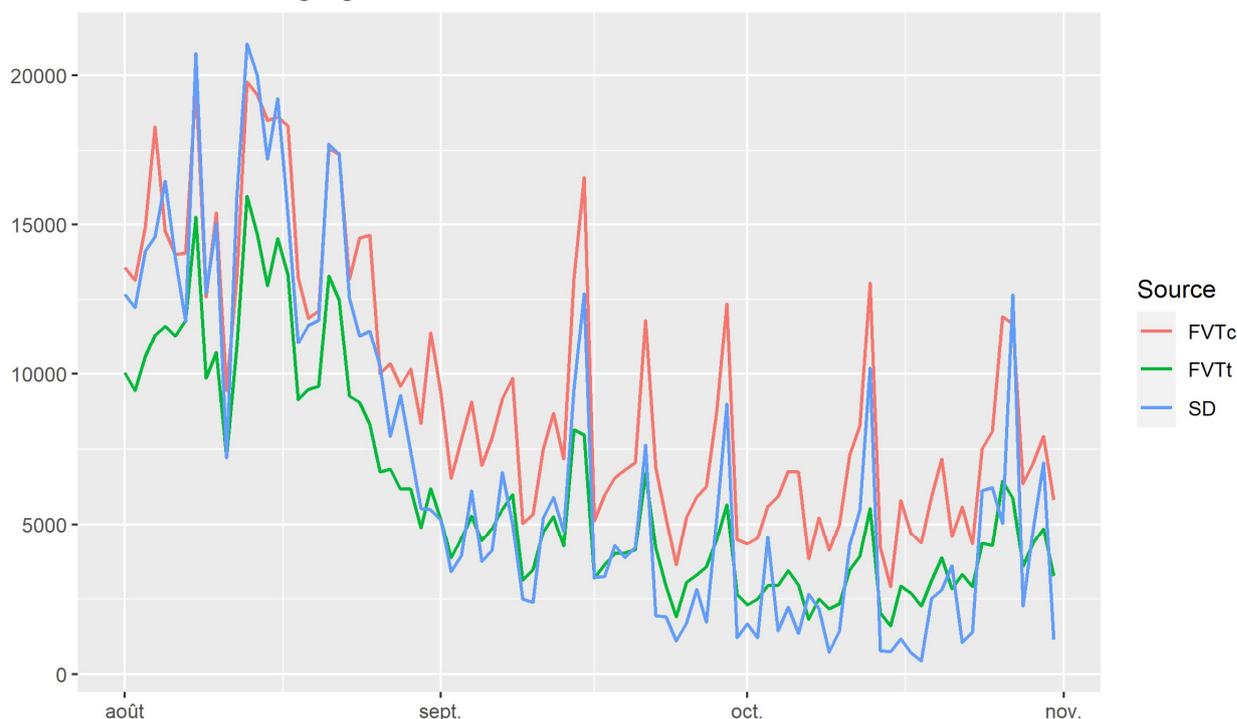


Figure 71 - Série Flux Vision Tourisme corrigée (FVTc), composante Touristes seule (FVTt) et agrégat des sites dénombrables (SD) sur la période août-octobre 2019

## 4.2. AUTRES DONNEES FLUX VISION

### 4.2.1. COMPARAISON AVEC LES DONNEES D'ENQUETE

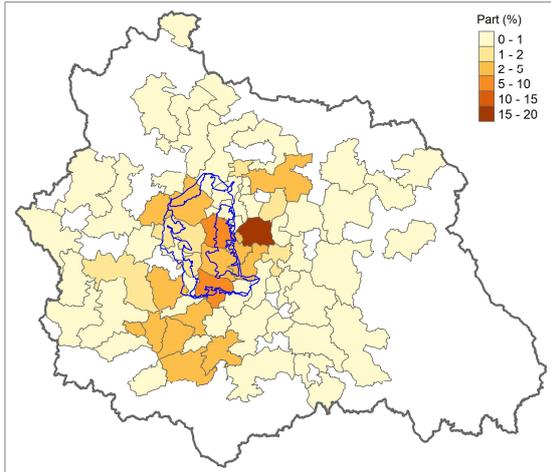
#### LIEUX DE NUITEE

Il est possible également de comparer les données Flux Vision portant sur la qualification des visiteurs avec les données de l'enquête visiteur Valo-Sites, réalisée entre mai 2019 et février 2020.

On peut par exemple regarder la répartition des lieux de nuitées des touristes présents sur la CDPFL en août 2019 (Figure 72). Les données ne sont toutefois pas tout à fait directement comparables, car il s'agit de proportion de nuitées (par groupement de commune) pour Flux Vision, et des proportions de touristes sur l'échantillon redressé (par commune) pour l'enquête.

On peut constater, dans les deux cas, une prépondérance de Clermont-Ferrand, ainsi qu'une importance marquée des communes de la CDPFL et du Sancy, plus nette encore sur les données FVT.

FVT - Répartition des nuitées (touristes - Août 2019)



GE - Répartition des nuitées (touristes - Août 2019)

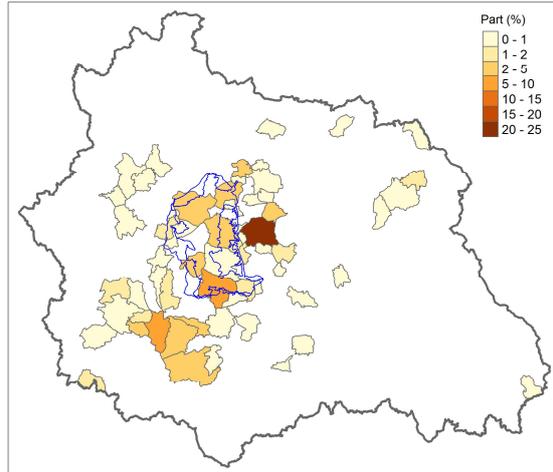


Figure 72 - Répartition des nuitées de touristes en août 2019 d'après les données Flux Vision Tourisme (à gauche) et d'après les données de l'enquête visiteur (à droite)

## CLASSES D'ÂGE

Sur les distributions en classes d'âge des visiteurs (Figure 73), sur la période d'août à octobre, on observe une part nettement plus importante des moins de 18 ans dans l'enquête, et une part moindre des plus âgés. Cela pourrait en partie s'expliquer par un échantillonnage plus marqué sur les périodes d'affluence, notamment les périodes de vacances, alors que l'échantillonnage Flux Vision est homogène sur toute la période, notamment en septembre-octobre, où il y a moins de familles et plus de retraités.

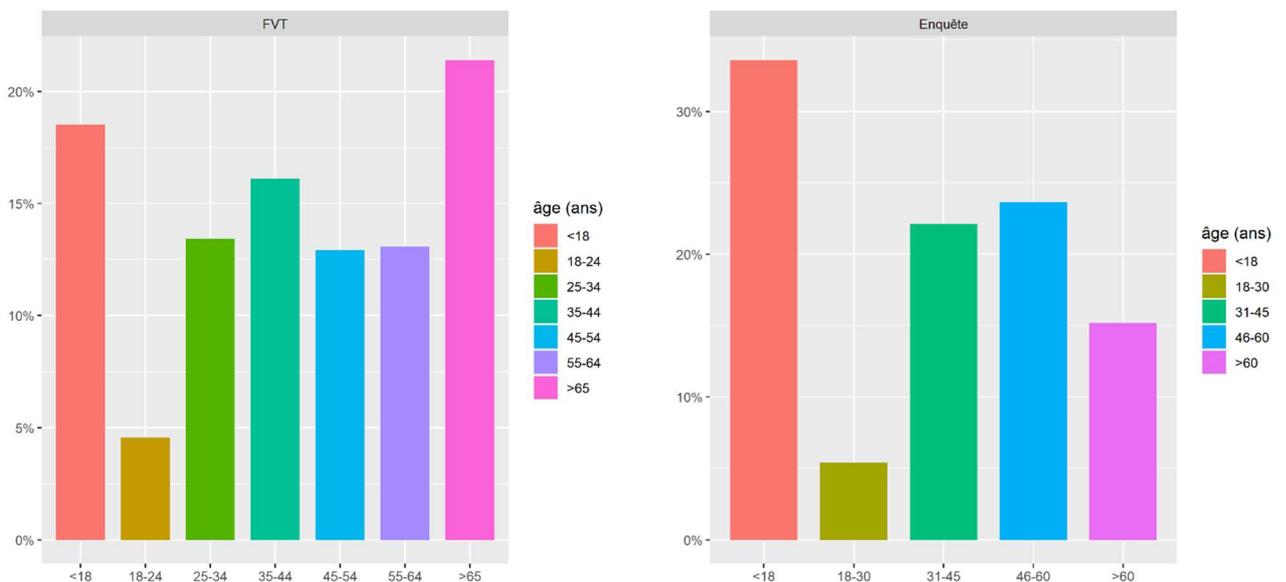


Figure 73 - Répartition en classes d'âge des visiteurs de la zone CDPFL sur la période août-octobre 2019 selon FVT et selon l'enquête 2019

En regardant, sur les données Flux Vision, l'évolution des proportions journalières de touristes par classe d'âge (Figure 74), on peut en effet constater que les moins de 18 ans sont plus présents en août (env. 20%), fluctuent entre 10 et 15% en septembre-octobre, et remontent un peu lors des vacances de la Toussaint. Inversement, les plus de 65 ans passent de 15% en août à près de 35% en septembre.

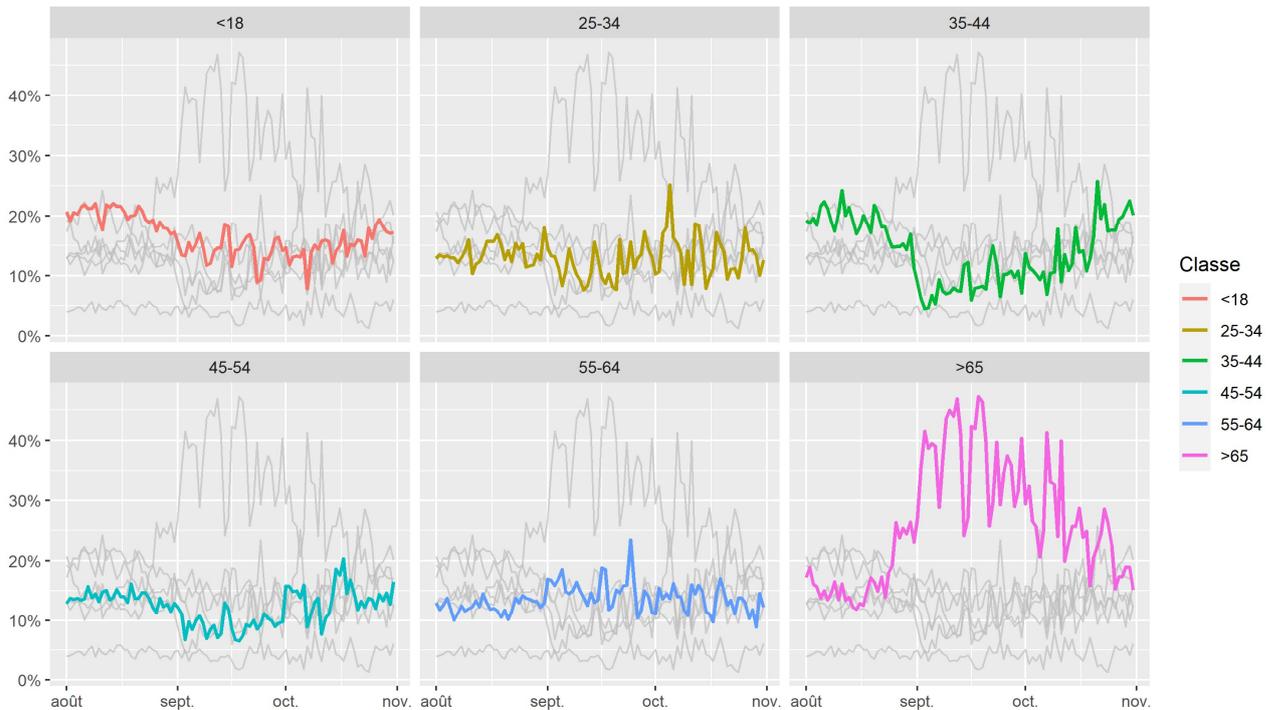


Figure 74 - Évolution des proportions de visiteurs par classe d'âge sur la période d'août à octobre 2019

#### 4.2.2. AUTRES INFORMATIONS FLUX VISION

##### CATEGORIES SOCIO-PROFESSIONNELLES

Sur la fréquentation par catégorie socio-professionnelle, en utilisant les données ventilées selon la typologie GéoLife d'Orange (Figure 75 et Figure 76), les évolutions des proportions par type, au cours de la période d'août à octobre 2019 (Figure 77), sont en général moins marquées que pour les classes d'âge. Cependant, on peut constater l'augmentation forte du « rural traditionnel » entre août (10%) et septembre (20%), jusqu'à près de 25% début octobre, suivi d'une baisse à 15% pendant les vacances de la Toussaint. En parallèle, on peut également noter une augmentation en septembre-octobre du type « résidences secondaires ». Inversement, les parts des types « pavillonnaire familial aisé » et « urbain familial aisé » augmentent sensiblement pendant les vacances scolaires.

**GéoLife se fait sur une base de 23 indicateurs pertinents pour la création des segments :**



Figure 75 - Indicateurs utilisés dans la typologie GéoLife de Flux Vision (source : Orange)

**GéoLife se décompose en 4 types de quartiers, 11 segments et 34 sous-segments**

Type de quartiers	Segments	Sous-Segmentation
urbain	<b>urbain dynamique</b>	<b>gentrification, jeune actif, étudiant &amp; cadre</b>
	<b>urbain familial aisé</b>	<b>privilegié, quartier historique, mix pavillon</b>
	<b>urbain classe moyenne</b>	<b>étudiant, cosmo, intermédiaire, sédentaire</b>
	<b>populaire</b>	<b>cœur historique, mix social, social</b>
	<b>urbain défavorisé</b>	<b>zus, barre sociale &amp; pavillon, centre ville paupérisé</b>
péri	<b>périurbain en croissance</b>	<b>quartier récent, multiculturel, résidentiel</b>
	<b>pavillonnaire familial aisé</b>	<b>favorisé, actif, sénior</b>
rural	<b>rural dynamique</b>	<b>rurbanisation, attractif, famille active</b>
	<b>rural ouvrier</b>	<b>famille, précarité, en déclin</b>
	<b>rural traditionnel</b>	<b>agricole, isolé, retraité</b>
rs	<b>résidence secondaire</b>	<b>saisonnier, mixte, tourisme vert</b>

Figure 76 - Typologie GéoLife de Flux Vision (source : Orange)

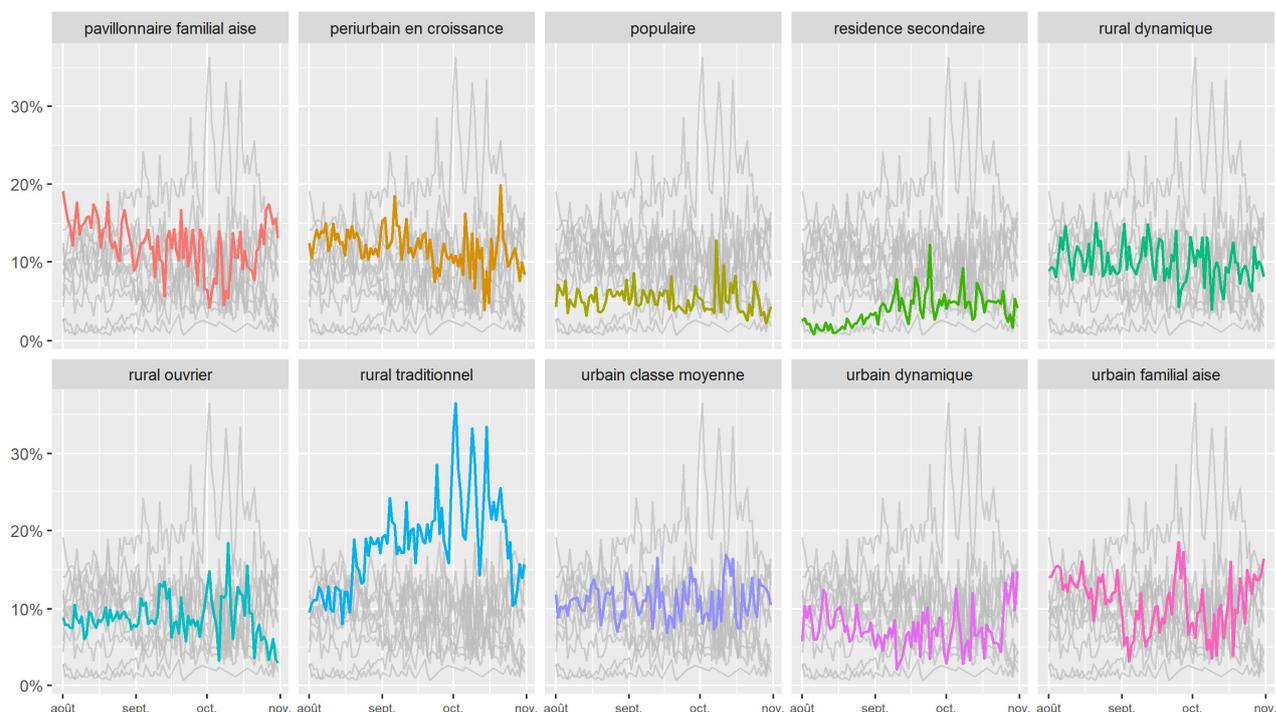


Figure 77 - Évolution des proportions de visiteurs par type GéoLife sur la période d'août à octobre 2019

## DUREES DE SEJOUR ET RECURRENCES

Certains indicateurs Flux Vision sont clairement inexploitable. C'est le cas notamment des durées de séjour. En effet, des arrivées et départs inopinés sont constatés, sans doute en lien avec les problèmes inhérents à la zone d'étude (relief, couverture). Il semble malgré tout possible, bien que plus difficilement, d'estimer des durées de séjour (Figure 78) et des nombres moyens de visites sur la zone d'étude (Figure 79), en exploitant les données de récurrence (il faut pour cela « bricoler » un peu les données Flux Vision).

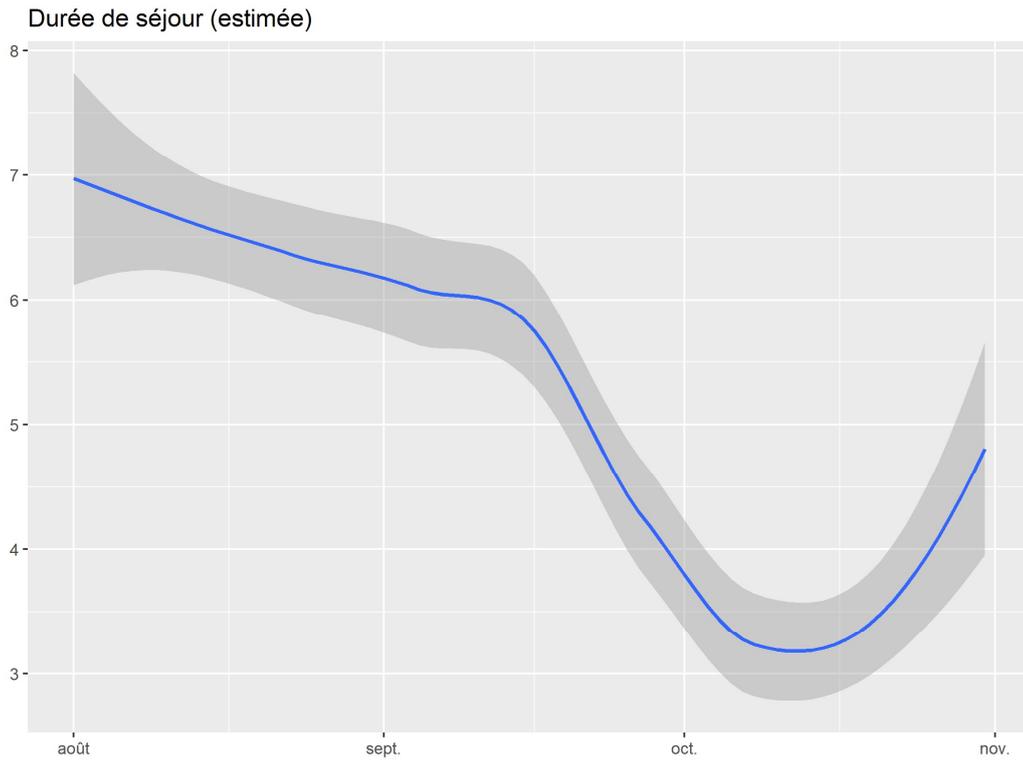


Figure 78 - Courbe lissée de l'évolution de la durée de séjour moyenne estimée à partir des données de récurrence Flux Vision (période août-octobre 2019)

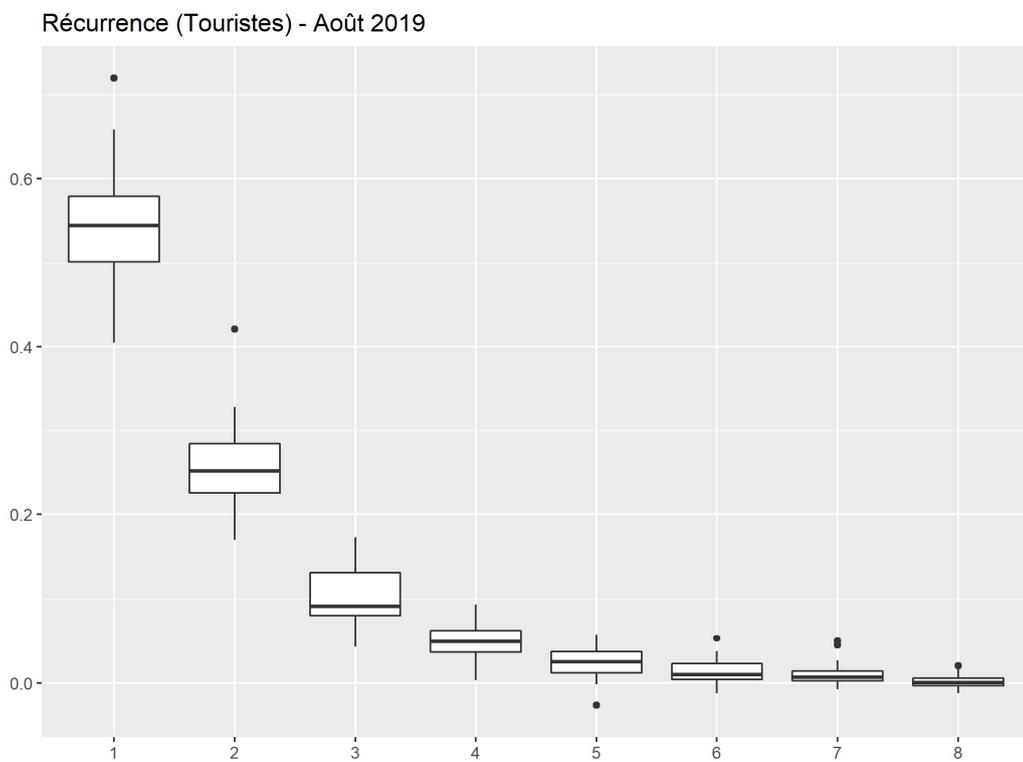


Figure 79 - Distribution des récurrences des touristes en août 2019 (nombre de jours passés sur CDPFL)

---

#### 4.2.3. SYNTHÈSE DES RESULTATS SUR FLUX VISION

L'utilisation de Flux Vision sur des zones telles que CDPFL est a priori problématique, en raison du relief, de la faible couverture mobile (nombre et génération des antennes), ainsi que la proximité de Clermont-Ferrand (problème de débords). Cela se traduit par diverses anomalies détectables dans les données (part excessive de résidents en transit, entrées-sorties...). De plus, une étude centrée sur une petite zone comme CDPFL nécessite une re-segmentation, car les touristes ne sont pas uniquement ceux logés dans la zone. Or, cette re-segmentation fondée sur le département de facturation n'est pas parfaitement équivalente. Enfin, les zones d'étude sont difficiles à établir, obligeant à trouver un compromis satisfaisant entre les manques et les risques de débords.

Cependant les résultats sur la zone de test (sous-zone centrée sur Vulcania, sur août 2019) ont montré que les données de Flux Vision étaient assez proches de l'assemblage des données de Lemptégy, Vulcania et des puits à proximité, avec une bonne corrélation entre les deux séries.

De même, sur la zone complète, les données sont proches et bien corrélées pour le mois d'août. En revanche, sur septembre-octobre, les estimations de Flux Vision, bien que corrélées, sont nettement plus élevées. Si la composante Touristes paraît crédible, car proche et très bien corrélée à l'agrégat des sites dénombrables, la composante Excursionnistes paraît trop élevée, même après correction. Elle pourrait s'expliquer en partie par des pratiques diffuses de locaux, sur les périodes hors vacances scolaire, par exemple des promeneurs et des sportifs utilisant les circuits ne montant pas aux puits et non-équipés d'éco-compteurs. Dans ce cas, en outre, l'extrapolation à partir des données d'éco-compteurs, appliquée uniformément sur l'année, pourrait accentuer encore cela, en minimisant trop les écarts au mois d'août et en exagérant ceux des mois de septembre et octobre.

Enfin, les données Flux Vision permettent de segmenter les visiteurs et apportent beaucoup d'informations (en plus des nombres de visiteurs), a priori exploitables pour les touristes, pour décrire les composantes de la fréquentation (données socio-démographiques, origines) et les types de séjour (lieu de nuitée, temps de présence, récurrence, durée de séjour, voire mobilité).

## 5. CONCLUSION

Cette étude a nécessité l'utilisation d'une grande quantité de données, de natures très diverses, nécessitant d'explorer finement les données de chaque source, de se familiariser avec les méthodes d'acquisition, d'identifier et de trouver des explications aux données atypiques. De plus, il a fallu rechercher des données complémentaires, souvent difficiles à obtenir. De nombreux échanges avec les acteurs concernés ont été nécessaires. Ce fut particulièrement le cas pour la mise en place des expérimentations visant à tester des approches alternatives de suivi de la fréquentation : sur le comptage routier d'une part, mais surtout pour le test de Flux Vision, pour lequel de nombreuses réunions avant la mise en place, mais également après, ainsi que de nombreux échanges ont été nécessaires tout au long du projet.

Une difficulté majeure de l'exercice est l'absence d'étalon pour juger de la fiabilité des estimations globales. Il est seulement possible de tester la cohérence des données, en comparant les résultats obtenus par les différentes approches, pour se faire une idée de la fiabilité de chacune, compte tenu de ses spécificités et limites propres. Les données de fréquentation obtenues à partir de ces approches alternatives, produites indépendamment des sources actuellement disponibles, pourraient apporter des éléments pour juger de la crédibilité des premières estimations.

Une autre difficulté provient de l'ouverture et de la complexité de la zone d'étude : entrées multiples, mobilités à travers la zone, transit, nombreux accès aux sentiers, nombreux sentiers, souvent interconnectés...). De plus, comme on l'a vu, tous les types de fréquentation sont présents sur la zone CDPFL, notamment certaines pratiques beaucoup plus diffuses (ex. promeneurs, sportifs sur des sentiers secondaires), et donc très difficiles à appréhender.

Si une approche telle que Flux Vision permettrait a priori d'appréhender cette fraction diffuse de la fréquentation, d'autres problèmes se posent, notamment liés aux questions de débords, ou encore à la segmentation des visiteurs. En revanche, cette approche permet d'estimer un nombre de visiteurs-jour, et n'est pas affectée a priori par le problème de double-compte, difficile à appréhender pour les estimations à partir des sites dénombrables.

D'autre part, les données disponibles sont très hétérogènes (cumuls annuels d'entrées pour les sites payants et données infra-horaires pour les éco-compteurs), mais aussi parcellaires (sentiers et sites non équipés d'éco-compteurs), ce qui ne permet pas de produire une estimation exhaustive fiable, et limite les possibilités de suivre la fréquentation plus finement dans le temps.

Enfin, les conversions des données de passage en données de visite, puis en visiteurs, sont très compliquées et nécessitent des informations complémentaires sur la fréquentation, difficilement estimables, même par enquête.

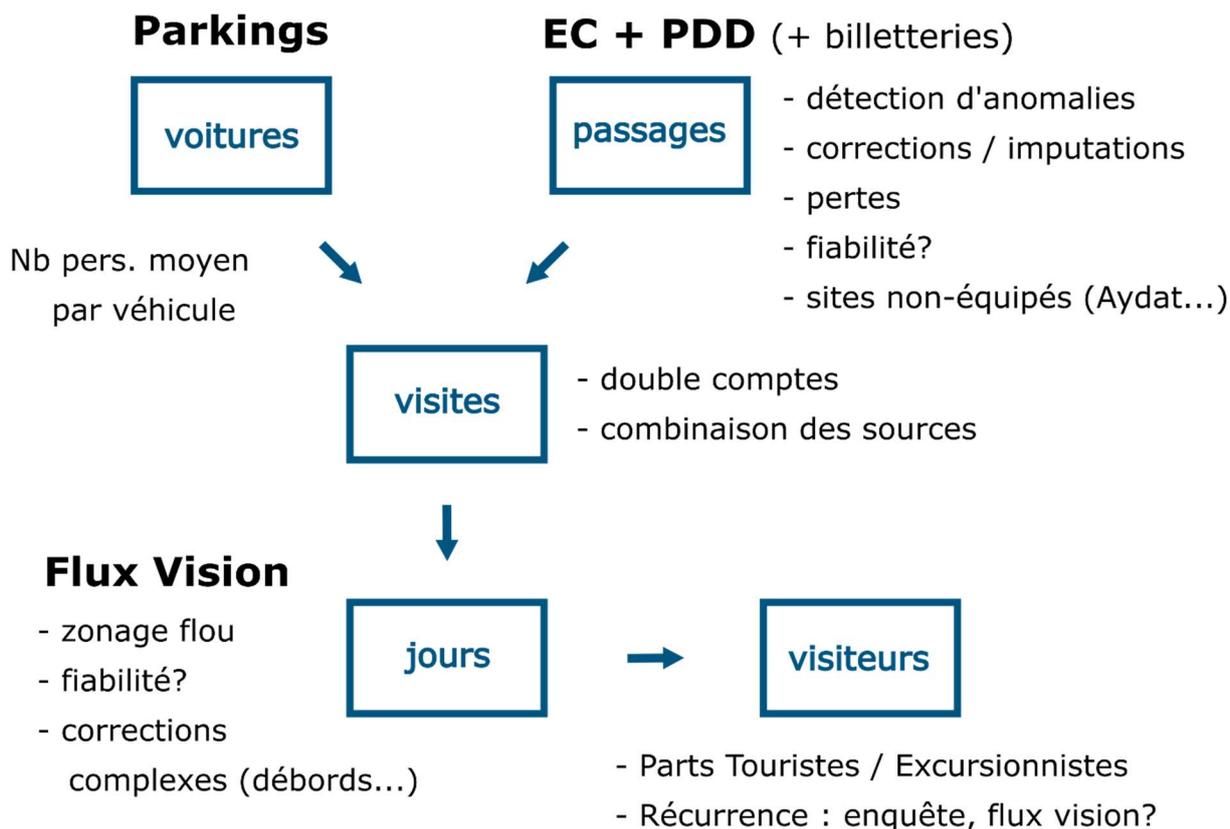


Figure 80 - Schéma intégrant les différentes sources de données et les types de données de fréquentation

## 5.1. AVANTAGES / INCONVENIENTS DES DIFFERENTES APPROCHES

On peut synthétiser les différents constats en termes d'avantages et inconvénients des différentes approches (Tableau 14). Globalement, l'utilisation des données disponibles (éco-compteurs et billetteries) permet d'établir une estimation plus ciblée, mais probablement non exhaustive (fréquentation diffuse), avec le problème de doubles comptes nécessitant des informations fiables sur les pratiques des visiteurs, tandis que l'approche Flux Vision, moins spécifique mais plus exhaustive, produit, en plus d'une estimation par composante (touristes, excursionnistes...), un contenu déjà très riche en terme de qualification de la fréquentation. Toutefois, la fiabilité de Flux Vision mériterait d'être mieux établie, avec des données complémentaires. Enfin, les comptages routiers semblent beaucoup moins opérationnels dans un contexte tel que la CDPFL, sans doute mieux adaptés à des comptages ponctuels, et à des sites avec parkings en bout de route.

Tableau 14 - Avantages et inconvénients des différentes approches testées

Approche	Avantages	Inconvénients
<b>Agrégation des sites dénombrables</b> (éco-compteurs + billetteries)	<ul style="list-style-type: none"> <li>- Données par site</li> <li>- Spécifique (tourisme/loisirs)</li> <li>- Reproductible</li> <li>- Coût limité</li> </ul>	<ul style="list-style-type: none"> <li>- Problème de double comptes</li> <li>- Peu exhaustif (sites/sentiers non équipés, tourisme diffus...)</li> <li>- Fiabilité (EC) ?</li> <li>- Perte de données (EC)</li> <li>- Pas d'info sur les visiteurs</li> </ul>
<b>Comptages routiers</b>	<ul style="list-style-type: none"> <li>- Complémentaire</li> <li>- Info / véhicule</li> </ul>	<ul style="list-style-type: none"> <li>- Mêmes inconvénients</li> <li>- Extrapolation difficile (fonction du nb de pers. par véhicule + autres transports)</li> <li>- Coût</li> </ul>
<b>Flux Vision</b>	<ul style="list-style-type: none"> <li>- « Exhaustif » (après redressement)</li> <li>- Reproductible (a priori)</li> <li>- Segmentation des visiteurs</li> <li>- Information riche</li> </ul>	<ul style="list-style-type: none"> <li>- Peu spécifique</li> <li>- Boîte noire</li> <li>- Fiabilité? (Touristes crédibles, Excursionnistes moins)</li> <li>- Coût (mais limité / enquête)</li> </ul>

## 5.2. CRITERES DE CHOIX

Dans tous les cas, le choix doit reposer sur une définition, ou une redéfinition, claire de l'objet du suivi selon les objectifs prioritaires :

- Suivre la fréquentation totale (touristes et excursionnistes) ou seulement les touristes ?
- Connaître la fréquentation toute l'année ou seulement sur la haute saison touristique ?
- Connaître la fréquentation sur quelques sites pour calculer des tendances, ou disposer de chiffres sur la fréquentation totale ?
- Observatoire du tourisme (impacts économiques) ou suivi de la fréquentation totale, pour mesurer des impacts sociaux, environnementaux ?
- Sur quelle zone exactement porte le suivi : Bien Unesco, Bien et zone tampon, ou plus large ? Avec ou sans Aydat ?
- Estimation globale ou par site ?
- Nombre de visites ou nombre de visiteurs ?
- Intérêt pour les mobilités entre zones (ex. Clermont-Ferrand, Sancy) ?

En outre, plusieurs autres critères peuvent entrer en ligne de compte, tels que :

- L'accès à des informations permettant de calibrer, planifier et redresser des enquêtes menées en parallèle ;
- L'accès plus ou moins rapide aux données de fréquentation.

### 5.3. PERSPECTIVES

La possibilité d'obtenir l'ensemble des données journalières de Vulcania sur le mois d'août 2019 (ainsi que les origines) permettrait enfin de trancher sur la fiabilité des estimations Flux Vision.

Globalement, disposer de données plus fines, notamment sur les données de billetterie des sites payants, offrirait d'autre possibilité dans le suivi, notamment pour analyser les évolutions saisonnières, voire pour modéliser la fréquentation.

Pour améliorer le suivi de la fréquentation sur Aydat, il serait intéressant de concevoir un protocole de suivi, même ponctuel, afin d'avoir a minima des éléments de calage, pour raffiner l'estimation.

En outre, des dispositifs, permanents ou ponctuels, pour estimer la fréquentation de sites non-équipés (éco-compteurs sur circuits, voire comptages vidéo sur parking) pourraient apporter des éclairages sur ces angles morts.

Enfin, seulement 3 mois de données de Flux Vision ont pu être analysés, mois peu représentatifs de l'ensemble des mois de l'année, ce qui rend hasardeux une extrapolation sur un an de l'estimation Flux Vision, et ne permet pas de tester la corrélation entre les sources sur différentes saisons. Dans le cas où les données Flux Vision pourraient être mieux validées, il serait intéressant de disposer d'une année complète, avec un choix de zones bien ciblé.

Si ces données sont jugées s'avèrent suffisamment fiables, elles pourraient être envisagées par exemple pour calibrer et/ou redresser une enquête auto-administrée en ligne, dans le cadre d'un suivi régulier de la fréquentation.

## BIBLIOGRAPHIE

- [1] J.-B. Marsat, K. B. Djongon, J. Veslot, et E. Polge, « La fréquentation touristique et ses effets économiques : des enjeux pour des politiques de gestion et de valorisation de sites naturels patrimoniaux. Application au site de la Chaînes des Puys - Faille de Limagne », in *16èmes Journées de Recherches en Sciences Sociales*, Clermont-ferrand, France, déc. 2022. [En ligne]. Disponible sur : <https://hal.science/hal-04210224>
- [2] V. Vlès, « De la difficulté de concilier développement et environnement », *Espaces tourisme & loisirs : revue mensuelle de réflexion du tourisme et des loisirs-Espaces : tourisme et loisirs-Espaces: tourisme, loisirs, environnement*, 2021, [En ligne]. Disponible sur : <https://hal.science/hal-03474733/>
- [3] Agence d'Urbanisme de la Région Nantaise, « Dénombrer les touristes : pas si simple... ! » [En ligne]. Disponible sur : <https://www.auran.org/articles/denombrent-les-touristes-pas-si-simple>
- [4] GIP Littoral aquitain, « Fréquentation des plages - Mesure de la fréquentation et modèle prédictif », 2018.
- [5] T. Lefebvre et S. Moncorps, « Les espaces protégés français : une pluralité d'outils au service de la conservation de la biodiversité », 2010.
- [6] Traces TPI, « Etude de fréquentation du Territoire du Grand Estérel », 2018. [En ligne]. Disponible sur : <https://pro.esterel-cotedazur.com/enquete-frequentation-massif-esterel-resultats/>
- [7] PN des Écrins, « Enquête de fréquentation touristique du Parc national des Ecrins », 2011. [En ligne]. Disponible sur : [https://www.ecrins-parcnational.fr/sites/ecrins-parcnational.com/files/fiche\\_doc/10550/2011-frequentation-ecrins-synthese-finale.pdf](https://www.ecrins-parcnational.fr/sites/ecrins-parcnational.com/files/fiche_doc/10550/2011-frequentation-ecrins-synthese-finale.pdf)
- [8] O. Boucher, « Analyse de la fréquentation touristique en lien avec les politiques de gestion et de valorisation du site « Chaîne des Puys - Faille de Limagne » inscrit comme bien naturel au patrimoine mondial de l'UNESCO / Elaborer un dispositif simplifié de suivi de la fréquentation sur le territoire de la Chaîne des Puys - Faille de Limagne », Rapport de stage de fin d'étude, 2021.
- [9] J.-B. Marsat, « Étude de la fréquentation de tourisme et loisirs du site Chaîne des Puys-Faille de Limagne : Nature de la fréquentation et effets économiques », rapport, 2022.
- [10] INSEE, « Données INSEE Filosofi 2015 carroyées 200m Auvergne-Rhône-Alpes - Data Europa EU ». [En ligne]. Disponible sur : <https://data.europa.eu/data/datasets/b10b9a4c-65a8-45a1-bd48-a476406eea2a?locale=fr>
- [11] INSEE, « Documentation des données carroyées Filosofi 2015 », juin 2019. [En ligne]. Disponible sur : <https://www.insee.fr/fr/statistiques/4176290?sommaire=4176305>
- [12] R. Rupf, M. Wernli, et R. Haller, « How to elaborate precise visitor numbers », in *Management for protection and sustainable development. Proceedings of the Fourth International Conference on Monitoring and Management of Visitor Flows in Recreational and Protected Areas. Montecatini Terme, Italy*, 2008, p. 161-164.

- [13] CD63, « Guide pour les professionnels du tourisme - Grands Sites Patrimoniaux du Puy de Dôme », Conseil départemental 63, 2019.
- [14] A. Moreau, « Mise en place d'un observatoire de la fréquentation sur le secteur de la Chaîne des Puys/Faille de Limagne dans le cadre du projet d'inscription de cet ensemble au Patrimoine Mondial de l'UNESCO », Rapport de stage, 2015.
- [15] H. Harder, P. Bro, N. Tradisaukas, et T. A. S. Nielsen, « Tracking Visitors in Public Parks: Experiences with GPS in Denmark », *Research in Urbanism Series*, vol. 1, p. 63-74, 2008.
- [16] S. Villain et V. Brizon, « Flux Vision Tourisme. Tourisme & Territoires juillet Communiqué de presse », Dossier de presse, juill. 2016. [En ligne]. Disponible sur : <https://docplayer.fr/57303836-Dossier-de-presse-flux-vision-tourisme-tourisme-territoires-juillet-communique-de-presse.html>
- [17] I. Brémond, V. Brizon, et C. Berlingué, « Bouches-du-Rhône Tourisme ouvre les portes du big data d'Orange. Le programme Flux Vision », *Revue Espaces*, n° 316, janv. 2014, [En ligne]. Disponible sur : <https://www.tourisme-espaces.com/doc/8973.bouches-rhone-tourisme-ouvre-portes-big-data-orange-programme-flux-vision.html>
- [18] V. Vlès et S. Clarimont, « Impacts des mesures de préservation des sites naturels exceptionnels : rapport final de recherche », UMR CERTOP 5044 ; UMR PASSAGES 5319, Research Report, févr. 2017. [En ligne]. Disponible sur : <https://hal.archives-ouvertes.fr/hal-01500485>
- [19] V. Vlès, « Régulation et contrôle des flux touristiques dans les espaces naturels protégés pyrénéens », janvier 2019. [En ligne]. Disponible sur : <https://hal.archives-ouvertes.fr/hal-01968686>
- [20] E. Marcelpoil et L. Langlois, « Protection de l'environnement et développement touristique en station : du conflit à l'organisation des proximités », *Développement durable et territoires. Économie, géographie, politique, droit, sociologie*, n° Dossier 7, Art. n° Dossier 7, mai 2006, doi: 10.4000/developpementdurable.2660.
- [21] J.-C. Gidrol, « Saison touristique d'été 2019 - Insee Focus - 171 ». [En ligne]. Disponible sur : <https://www.insee.fr/fr/statistiques/4254937#documentation>
- [22] L. Dauphin, M.-A. Le Garrec, et F. Tardieu, « Les vacances des Français depuis 40 ans », *Le tourisme en France, édition*, vol. 2008, p. 31-40, 2008.
- [23] B. Kokou, J.-B. Marsat, E. Polge, et J. Veslot, « Contribution à l'évaluation des effets économiques de la fréquentation d'un grand site naturel protégé - Le cas de la Chaîne des Puys-Faille de Limagne, site inscrit en 2018 sur la liste UNESCO du Patrimoine mondial », in *Territoire (s) et numérique innovations, mutations et décisions*, 2021.
- [24] C. Berlingué et J.-M. Contet, « Flux Vision Tourisme. Le big data au service de l'observation des flux touristiques », *Revue Espaces*, vol. 335, mars 2017.
- [25] ATEN, « Etudier la fréquentation touristique dans les espaces naturels », A.T.E.N., 1999.
- [26] M. Carreno et M. Marchand, « The monitoring of movements: analysis method of spatial

- tourism. », *Cahiers Espaces*, n° 60, p. 32-40, 1999.
- [27] G. Cousin et F. Hillaireau, « Can Mobile Phone Data Improve the Measurement of International Tourism in France? », *Economie et Statistique*, vol. 505, n° 1, p. 89-107, 2018.
- [28] M. Carreno et F. Marchand, « L'observation par les flux. Outil de l'analyse touristique territoriale », *cahier Espaces*, no. 60. 1999. [En ligne]. Disponible sur : <https://www.tourisme-espaces.com/doc/1997.1-observation-flux-outil-analyse-touristique-territoriale.html>
- [29] SIVOM Grand Site Gorges du Tarn, de la Jonte, et Causses, « Note sur les systèmes de comptage automatique des personnes », 2017. [En ligne]. Disponible sur : <https://docplayer.fr/20166723-Note-sur-les-systemes-de-comptage-automatique-des-personnes.html>
- [30] L. Kajala, *Visitor monitoring in nature areas: a manual based on experiences from the Nordic and Baltic countries*. Stockholm: Swedish Environmental Protection Agency, 2007.
- [31] H. Job, L. Majewski, M. Engelbauer, S. Bittlingmaier, et M. Woltering, « Establishing a standard for park visitation analyses: Insights from Germany », *Journal of Outdoor Recreation and Tourism*, vol. 35, p. 100404, 2021.
- [32] A. Spenceley *et al.*, « Visitors count! Guidance for protected areas on the economic analysis of visitation », 2021.
- [33] J. Jurado Rota, M. Y. Pérez Albert, et D. Serrano Gine, « Visitor monitoring in protected areas: an approach to Natura 2000 sites using Volunteered Geographic Information (VGI) », *Geografisk Tidsskrift-Danish Journal of Geography*, vol. 119, n° 1, p. 69-83, 2019.
- [34] R. Bushell et T. Griffin, « Monitoring visitor experiences in protected areas », *Int. J. Protect. Area Manage*, vol. 16, n° 2, p. 25-33, 2006.
- [35] W. L. Hadwen, W. Hill, et C. M. Pickering, « Icons under threat: Why monitoring visitors and their ecological impacts in protected areas matters », *Ecological Management & Restoration*, vol. 8, n° 3, p. 177-181, 2007.
- [36] A. Muhar, A. Arnberger, et C. Brandenburg, « Methods for Visitor Monitoring in Recreational and Protected Areas: An Overview », p. 73, 2002.
- [37] C. Walden-Schreiner, S. D. Rossi, A. Barros, C. Pickering, et Y.-F. Leung, « Using crowd-sourced photos to assess seasonal patterns of visitor use in mountain-protected areas », *Ambio*, vol. 47, n° 7, p. 781-793, 2018.
- [38] H. Tenkanen *et al.*, « Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas », *Scientific reports*, vol. 7, n° 1, p. 1-11, 2017.
- [39] V. Heikinheimo, E. D. Minin, H. Tenkanen, A. Hausmann, J. Erkkonen, et T. Toivonen, « User-generated geographic information for visitor monitoring in a national park: A

comparison of social media data and visitor survey », *ISPRS International Journal of Geo-Information*, vol. 6, n° 3, p. 85, 2017.

- [40] P. S. Ziesler et D. Pettebone, « Counting on visitors: A review of methods and applications for the National Park Service's visitor use statistics program », *Journal of Park and Recreation Administration*, vol. 36, n° 1, 2018.
- [41] J. B. Read, M. J. Daniels, et L. K. Harmon, « Implementing technology-based visitor counts in parks: A methodological overview », *Journal of Park and Recreation Administration*, vol. 39, n° 1, 2021.
- [42] D. English et J. M. Bowker, « Introduction to the special issue on visitor monitoring », *Journal of Park and Recreation Administration*, vol. 36, n° 1, 2018.
- [43] T. Rudenko, « Visitor Monitoring in Protected Areas: Concepts, Methods and Trends », *InterConf*, p. 295-305, 2021.
- [44] J. Breyne, F. E. Montero de Oliviera, J. Abiltrup, et M. Dufrière, « Une estimation de la fréquentation touristique dans les espaces naturels de l'Ardenne en utilisant des technologies nouvelles », ULiege, GxABT, 2021.
- [45] N. Le Corre, S. L. Le Berre, L. Brigand, et I. Peuziat, « Comment étudier et suivre la fréquentation dans les espaces littoraux, marins et insulaires? De l'état de l'art à une vision prospective de la recherche », *EchoGéo*, n° 19, 2012.

## LISTE DES FIGURES

Figure 1 - Schéma général du projet VALO-SITES .....	10
Figure 2 - Carte de localisation du site Chaîne des Puys - Faille de Limagne .....	11
Figure 3 - Carte des différents éléments d'intérêt du bien inscrit, et périmètres des zones classées et protégées.....	11
Figure 4 - Carte des communes, incluses ou jouxtant la zone d'étude .....	12
Figure 5 - Carte des comptages routiers effectués sur la zone d'étude en 2019 .....	13
Figure 6 - Extrait de la carte touristique Auvergne-Rhône-Alpes.....	14
Figure 7 - Extrait de la carte scan 25 centrée sur la zone du Puy de Pariou.....	15
Figure 8 - Schéma du fonctionnement de l'éco-compteur .....	20
Figure 9 - Photos d'éco-compteur in situ.....	20
Figure 10 - Copie d'écran de l'interface ecovisio permettant la visualisation et l'extraction des données d'éco-compteur .....	21
Figure 11 - Description du capteur PYRO .....	22
Figure 12 - Répartition spatiale des éco-compteurs sur la zone d'étude.....	24
Figure 13 - Séries journalières par sens de passage, sur le mois de septembre 2019, pour les 16 éco-compteurs disponibles.....	25
Figure 14 - Proportion de données manquantes par éco-compteurs et par année .....	26
Figure 15 - Proportion de données manquantes par année pour 16 éco-compteurs de la zone d'étude (2016-2019).....	27
Figure 16 - Carte scan25 centrée sur les puys du nord de la zone d'étude .....	28
Figure 17 - Carte scan25 centrée sur la zone Pariou - Puy de Dôme .....	28
Figure 18 - Format des données du Panoramique des Dômes.....	30
Figure 19 - Évolution du nombre annuel d'entrées sur les différents sites payants de la zone d'étude entre 2010 et 2019 .....	31
Figure 20 - Évolution du nombre annuel d'entrées par site .....	32
Figure 21 - Calendrier d'ouverture des sites payants de CPFL .....	33
Figure 22 - Trails dans ou proches de la zone CPFL en 2019.....	36
Figure 23 - Baromètre de la fréquentation dans les hébergements collectifs en 2019 dans le Puy de Dôme – CRT Auvergne-Rhône-Alpes .....	37
Figure 24 - Module d'Evaluation du poids Touristique des Territoires (METT) - Zone UNESCO en mars 2021 .....	37
Figure 25 - Carte des éco-compteurs utilisés .....	38
Figure 26 - Séries brutes (2016-2019).....	40
Figure 27 - Détection des pics anormaux (triple filtre) .....	41
Figure 28 - Séries sans les pics anormaux.....	42
Figure 29 - Éco-compteur de Jumes (début 2016) .....	43
Figure 30 - Séries lissées (par année).....	44

Figure 31 - Détection des séquences anormalement basses (filtre médiane).....	45
Figure 32 - Répartition des séquences anormalement basses (bleu clair).....	46
Figure 33 - Séries IN et OUT de l'éco-compteur Pariou-Nord avec séquences anormalement basses.....	47
Figure 34 - Séries des moyennes mensuelles .....	48
Figure 35 - Indicateur d'inversion.....	49
Figure 36 - Séries IN et OUT de l'éco-compteur Pariou-Nord (par année).....	50
Figure 37 - Séries des écarts de cumul IN - OUT par site.....	51
Figure 38 - Séries brutes (bleu) et imputées (rouge) .....	53
Figure 39 - Fréquentation journalière d'Aydat (estimée) .....	57
Figure 40 - Série complète de la fréquentation de la CDPFL en 2019 obtenue par agrégation des séries des sites dénombrables.....	58
Figure 41 - Carte des parkings de la zone d'étude .....	62
Figure 42 - Dispositif de comptage mis en place dans le Parc National des Écrins sur un parking situé en bout de route .....	62
Figure 43 - Carte de la répartition des compteurs routiers posé durant l'été 2019 .....	64
Figure 44 - Photos des lieux de comptages .....	65
Figure 45 - Comptages réalisés par le Service des routes du CD63.....	66
Figure 46 - Comptages routiers en amont et aval du parking de Vache-Lassolas, avec dysfonctionnement du compteur Sud après le 18 juillet 2019 .....	67
Figure 47 - Flux entrant et sortant pour les compteurs amont et aval au niveau du parking de Lemptégy.....	67
Figure 48 - Comptage des véhicules sur le parking du Panoramique vs. cumul train + sentier des chèvres.....	68
Figure 49 - Photo du parking des Goules avec représentation schématique du positionnement des compteurs routiers en amont et en aval.....	69
Figure 50 - Schéma des principaux flux au niveau du parking .....	69
Figure 51 - Nombre de montées par l'éco-compteur des muletiers (rouge) et nombre de véhicules ayant stationné sur le parking du col de Ceysnat (x5) entre le 17 juillet et le 22 août .....	70
Figure 52 - Séquences anormalement basses en juillet sur les données de l'éco-compteur du sentier des muletiers .....	71
Figure 53 - Photo de l'entrée du parking du Volcan de Lemptégy avec les différentes voies (A) et schéma des flux mesurés par vidéo (B).....	72
Figure 54 - Séries des entrées et sorties du parking du Volcan de Lemptégy le 14/08/2019.....	73
Figure 55 - Schéma simplifié du fonctionnement de Flux Vision.....	75
Figure 56 – Étapes du redressement des données dans Flux Vision .....	76
Figure 57 - Segmentation des visiteurs dans Flux Vision Tourisme.....	77
Figure 58 - Problèmes de définition de la zone d'étude.....	79

Figure 59 - Exemple de visualisation des données via l'interface powerBI de Flux Vision (département du Cantal) .....	80
Figure 60 - Zone A (zone complète étendue).....	82
Figure 61 - Zone B (zone complète réduite) .....	82
Figure 62 - Zone C (sous-zone centrée sur Lemptégy et Vulcania).....	83
Figure 63 - Zone D (sous-zone centrée sur Pariou – Puy de Dôme).....	84
Figure 64 - Zones d'étude Flux Vision .....	88
Figure 65 - Carte des groupements de communes selon les valeurs du ratio des nuitées « Excursionnistes » sur la population.....	94
Figure 66 - Distribution des ratios Nuitées / Population, et choix d'un seuil .....	95
Figure 67 - Données journalières Flux Vision vs. Vulcania+Lemptégy (août 2019) ..	96
Figure 68 - Séries Flux Vision vs Agrégats billetteries + éco-compteurs.....	97
Figure 69 - Séries Flux Vision Tourisme corrigée et agrégat des sites dénombrables (sans Volvic) sur la ZCR au mois d'août 2019.....	98
Figure 70 - Série Flux Vision Tourisme corrigée (FVTc) et agrégat des sites dénombrables (SD) sur la période août-octobre 2019.....	99
Figure 71 - Série Flux Vision Tourisme corrigée (FVTc), composante Touristes seule (FVTt) et agrégat des sites dénombrables (SD) sur la période août-octobre 2019 .....	100
Figure 72 - Répartition des nuitées de touristes en août 2019 d'après les données Flux Vision Tourisme (à gauche) et d'après les données de l'enquête visiteur (à droite).....	101
Figure 73 - Répartition en classes d'âge des visiteurs de la zone CDPFL sur la période août-octobre 2019 selon FVT et selon l'enquête 2019 .....	101
Figure 74 - Évolution des proportions de visiteurs par classe d'âge sur la période d'août à octobre 2019 .....	102
Figure 75 - Indicateurs utilisés dans la typologie GéoLife de Flux Vision (source : Orange).....	103
Figure 76 - Typologie GéoLife de Flux Vision.....	103
Figure 77 - Évolution des proportions de visiteurs par type GéoLife sur la période d'août à octobre 2019 .....	104
Figure 78 - Courbe lissée de l'évolution de la durée de séjour moyenne estimée à partir des données de récurrence Flux Vision (période août-octobre 2019).....	105
Figure 79 - Distribution des récurrences des touristes en août 2019 (nombre de jours passés sur CDPFL).....	105
Figure 80 - Schéma intégrant les différentes sources de données et les types de données de fréquentation.....	108

## LISTE DES TABLEAUX

Tableau 1 - Différents modes d'acquisition des données de fréquentation .....	7
Tableau 2 - Méthodes de comptage.....	7
Tableau 3 - Différentes mesures de la fréquentation.....	9
Tableau 4 - Données disponibles pour les sites dénombrables .....	19
Tableau 5 - Coefficients multiplicatifs permettant d'extrapoler le nombre réel de visiteurs sur ou à proximité du puy .....	29
Tableau 6 - Liste des parkings de la zone d'étude .....	61
Tableau 7 - Segmentation des "dormants" .....	77
Tableau 8 - Segmentation des "non dormants" .....	78
Tableau 9 - Liste des fichiers de données de l'offre Flux Vision Tourisme .....	86
Tableau 10 - Liste des fichiers de données de l'offre Flux Vision Géomarketing .....	87
Tableau 11 - Résidents Flux Vision et population INSEE sur les zones Flux Vision..	89
Tableau 12 - Regroupements de communes Flux Vision sur le Puy de Dôme avec ratios de nuitées sur population.....	90
Tableau 13 - Proportion de visiteurs par origine selon différentes sources .....	98
Tableau 14 - Avantages et inconvénients des différentes approches testées.....	109