



HAL
open science

Étude de la fréquentation de tourisme et loisirs du site Chaîne des Puys-Faille de Limagne. Tâches T1B et T4A du projet Valo-sites : Nature de la fréquentation et effets économiques.

Jean-Bernard Marsat, Kokou Boris Djongon, Claire Bussière, Nadia Guiffant,
Jacques Veslot, Marie Batifol, Orane Boucher, Alexia Gignon, Etienne Polge

► To cite this version:

Jean-Bernard Marsat, Kokou Boris Djongon, Claire Bussière, Nadia Guiffant, Jacques Veslot, et al..
Étude de la fréquentation de tourisme et loisirs du site Chaîne des Puys-Faille de Limagne. Tâches
T1B et T4A du projet Valo-sites : Nature de la fréquentation et effets économiques.. INRAE. 2022.
hal-04466524

HAL Id: hal-04466524

<https://hal.science/hal-04466524v1>

Submitted on 22 Feb 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.

Projet Valo-Sites

Méthode d'évaluation des effets socio-économiques des politiques
de gestion et valorisation de sites naturels

Etude de la fréquentation de tourisme et loisirs du site Chaîne des Puys-Faille de Limagne

Tâches T1-B et T4-1 :

Nature de la fréquentation et effets économiques

Partie I : Cadre, méthodes, estimation économique

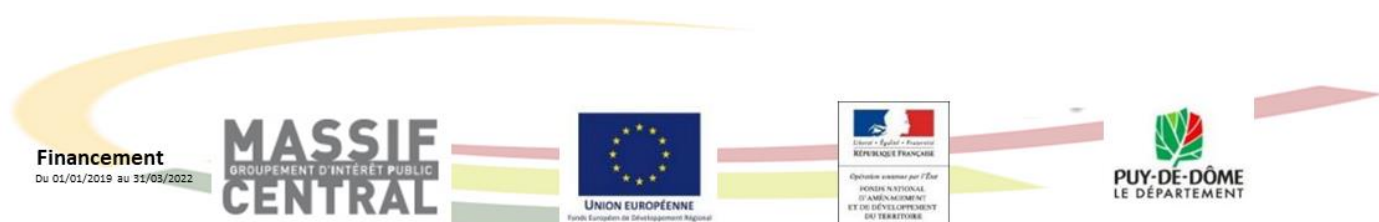
Pré-rapport final (mars 2022)

Auteurs :

Jean-Bernard Marsat et Kokou Boris Djongon

Avec la participation de : Claire Bussière, Nadia Guiffant, Etienne Polge, Jacques Veslot

Marie Batifol, Orane Boucher, Alexia Gignon



Ce document est un « *pré-rapport final* » relatif à une partie des objectifs et des tâches du projet Valo-Sites.

Il sera complété dans sa version définitive, ne serait-ce que sur les aspects de mise en forme et de documentation des tableaux, des figures, de la bibliographie et des annexes.

Table des matières

| | | |
|--------|---|----|
| I. | Introduction | 1 |
| I.1 | Le contexte | 1 |
| I.2 | L'organisation générale des tâches du projet | 3 |
| I.3 | Les objectifs des tâches T1-B et T4-1 | 4 |
| II. | Problématiques, cadres théoriques et méthodologiques..... | 5 |
| II.1 | Problématique générale : valeurs d'un espace naturel, politiques de gestion et valorisation, évaluations diverses | 5 |
| II.1.1 | Distinguer état (valeurs) de l'espace, évolution de cet état, et évaluation de politiques menées | 5 |
| II.1.2 | Distinguer les effets propres, distinguer gestion et valorisation..... | 7 |
| II.1.3 | Valeurs d'un espace naturel..... | 8 |
| II.2 | La valorisation économique des espaces <i>protégés</i> | 10 |
| II.3 | La part importante de la fréquentation de tourisme et loisirs dans la valorisation économique (et sociale) des espaces naturels | 10 |
| II.4 | Les dimensions sociales et économiques de la fréquentation d'un site naturel (protégé) | 11 |
| II.4.1 | Enjeux | 11 |
| II.4.2 | Dimensions marchandes et non marchandes de la valorisation..... | 12 |
| II.4.3 | La relation du touriste/excursionniste à l'espace naturel : directe ou via le « triangle des services » | 12 |
| II.4.4 | Géographie du phénomène de tourisme et loisirs liés à la visite d'un site..... | 15 |
| II.5 | La connaissance de la fréquentation des espaces naturels (protégés) : objectifs, dispositifs d'observation, définitions et notions-clés | 19 |
| II.5.1 | Objectifs..... | 19 |
| II.5.2 | Panorama d'études et de méthodes d'étude de la fréquentation..... | 19 |
| II.5.3 | Définitions et notions-clés | 21 |
| II.6 | Evaluation des retombées économiques..... | 23 |

| | | |
|---------|---|----|
| II.6.1 | Estimation des dépenses de la fréquentation d'un site : quels types et parts des dépenses à prendre en compte ? | 25 |
| II.6.2 | Mesure de l'activité économique..... | 30 |
| II.7 | Etude de la fréquentation d'un site : aspects méthodologiques détaillés | 37 |
| II.7.1 | Méthode d'échantillonnage | 37 |
| II.7.2 | Traitement des non-réponses et redressement d'échantillon | 43 |
| II.7.3 | Traitement des voyages et séjours à forfait | 46 |
| III. | Le cas de la CPFL : l'enquête socio-économique auprès des visiteurs..... | 47 |
| III.1 | La Chaîne des Puys – faille de Limagne | 47 |
| III.2 | Conception et réalisation de l'enquête | 47 |
| III.2.1 | Le type d'enquête | 47 |
| III.2.2 | Questionnement et questionnaire..... | 48 |
| III.2.3 | Protocole d'enquête | 51 |
| III.3 | Saisie et premiers traitements des données | 64 |
| III.3.1 | Saisie sur le terrain | 64 |
| III.3.2 | Déchargement des données, ajout de données de traçage | 64 |
| III.3.3 | Statut des données | 64 |
| III.3.4 | conversions de données | 65 |
| III.3.5 | Vérifications / corrections | 66 |
| III.3.6 | Imputations | 68 |
| III.3.7 | Simulations de compléments ou raffinements éventuels..... | 69 |
| III.3.8 | Comparaisons déclaré/imputé, corrections..... | 71 |
| III.3.9 | Traçabilité des opérations | 72 |
| III.4 | Redressement des données collectées | 73 |
| III.4.1 | Pondération probabiliste ex post..... | 73 |
| III.4.2 | Post-stratification par période ou par lieu-période | 74 |

| | | |
|---------|---|----|
| III.4.3 | Qualités et limites de l'enquête et des traitements de données, précision des résultats..... | 77 |
| IV. | CPFL : résultats de l'enquête, partie I : éléments nécessaires au calcul d'effet économique | 80 |
| IV.1 | Les motivations pour la CPFL : calcul du coefficient de « l'effet CPFL ».... | 80 |
| IV.1.1 | Les motivations des excursionnistes : hypothèse effet CPFL 100% | 80 |
| IV.1.2 | Les motivations des touristes : calcul de « l'effet CPFL » (touristes)..... | 80 |
| IV.1.3 | Question principale..... | 81 |
| IV.1.4 | Question de contrôle | 82 |
| IV.1.5 | Modulation de C1 : coefficient C2 | 84 |
| IV.2 | Les activités dénombrables des visiteurs : contribution à la mesure de la fréquentation | 86 |
| IV.2.1 | Principe de l'estimation d'une fréquentation-plancher de la CPFL en visiteurs (et non en visites), à partir du nombre de visites..... | 86 |
| IV.2.2 | Calcul des taux de comptes multiples..... | 87 |
| IV.2.3 | Le cas des visiteurs non dénombrables..... | 88 |
| IV.2.4 | Le cas des « excursionnistes répétés »..... | 89 |
| IV.2.5 | Résultat : estimation d'une fréquentation-plancher | 90 |
| IV.3 | Les dépenses des visiteurs : premier aperçu..... | 93 |
| IV.3.1 | Hierarchie des dépenses..... | 93 |
| IV.3.2 | Comparaison avec les données du SDT (Suivi de la Demande Touristique, INSEE-Kantar) | 94 |
| V. | CPFL : mesure de l'activité économique générée par la visite du site..... | 96 |
| V.1 | Détermination des impacts directs | 97 |
| V.1.1 | Localisation des dépenses..... | 97 |
| V.1.2 | Affectation de tout ou partie des dépenses des visiteurs à « l'effet CPFL » (rappel) | 99 |
| V.1.3 | Calcul des dépenses moyennes par visiteur | 99 |

| | | |
|--------|---|-----|
| V.1.4 | Calcul des dépenses des visiteurs dans chaque département : totaux, importances économiques, impacts économiques..... | 101 |
| V.1.5 | La conversion des dépenses directes en valeur ajoutée, emplois et de revenus. | 106 |
| V.2 | Estimation des effets induits par les effets directs | 123 |
| V.2.1 | Méthodologie..... | 123 |
| V.2.2 | Résultats..... | 126 |
| VI. | Bibliographie..... | 129 |
| VII. | Bibliographie complémentaire | 137 |
| VIII. | Annexes..... | 139 |
| VIII.1 | Annexe : Caractéristique des zones d'impact..... | 139 |
| VIII.2 | Annexe : Les traitements des données | 142 |
| VIII.3 | Annexe : Structure des cellules du fichier exporté LimeSurvey vers Excel | 147 |
| VIII.4 | Annexe : Quelques corrections/imputations effectuées dans le fichier..... | 148 |
| VIII.5 | Annexe : Précisions mathématiques : moyennes, variance pondérée | 150 |

I Acronymes & abréviations

| | |
|-----------------|---|
| <i>UPE</i> | Unités Primaires d'Echantillonnage |
| <i>MGPP</i> | Méthode Généralisée du Partage des Poids |
| <i>PDD</i> | Puys de Dôme |
| <i>INED</i> | Institut national d'études démographiques |
| <i>CPFL</i> | Chaîne des Puys Faille de Limagne |
| <i>INSEE</i> | Institut National de la Statistique et des Etudes Economiques |
| <i>TMA</i> | Trade Market Analysis |
| <i>EUROSTAT</i> | Direction Générale de la Commission Européenne chargée de l'information statistique |
| <i>USDA FS</i> | Service des forêts des États-Unis |

II Liste des tableaux

| | |
|---|-----|
| Tableau 1 : Le lieu d'étude suggéré en fonction du type de zone..... | 38 |
| Tableau 2 : Le modèle d'enquête proposé en fonction du type d'informations à collecter | 39 |
| Tableau 3 : Sources des données par site d'enquête | 56 |
| Tableau 4 : Nombre de visites par sites et par mois en 2019 | 57 |
| Tableau 5 : Localisation des dépenses | 98 |
| Tableau 6 : Dépenses moyennes par visiteur selon le type de visiteurs et le type de dépenses considérées (euros)..... | 100 |
| Tableau 7 : Dépenses moyennes par type de visiteurs et le type de dépenses considérées dans les autres départements (euros) | 100 |
| Tableau 8 : Fréquentation de la CPFL : détail pour niveaux-planchers de 800 ou 900 000 visiteurs | 101 |
| Tableau 9 : Dépenses totales selon le type de visiteurs et le type de dépenses considérées (euros)..... | 102 |
| Tableau 10 : Dépenses totales selon le type de visiteurs et le type de dépenses considérées dans les autres départements (euros) | 102 |
| Tableau 11 : Emploi total dans les départements en 2017 | 106 |
| Tableau 12 : Part et emploi salarié touristique | 107 |
| Tableau 13 : Répartition de l'emploi salarié touristique selon les activités | 108 |
| Tableau 14 : Ventilation par activité de l'emploi salarié touristique dans les départements . | 109 |
| Tableau 15 : Données nationales sur la structure de l'emploi (en millier) dans les branches « hébergement-restauration », « commerce de détail » et « autres services » | 109 |
| Tableau 16 : Estimation de l'emploi touristique total dans l'Allier en 2017 : répartition par activité | 109 |
| Tableau 17 : Estimation de l'emploi touristique total dans le Cantal en 2017 : répartition par activité | 110 |
| Tableau 18 : Estimation de l'emploi touristique total dans la Haute-Loire en 2017 : répartition par activité | 110 |
| Tableau 19 : Estimation de l'emploi touristique total dans la Puy-de-Dôme en 2017 : répartition par activité | 111 |
| Tableau 20 : Estimation de l'emploi touristique total par département en 2017 | 112 |

| | |
|--|-----|
| Tableau 21 : Données nationales sur les branches « hébergement-restauration », « Commerce de détail » et « autres services » | 113 |
| Tableau 22 : Ratios nationaux de branche (euros par emploi, 2017) | 113 |
| Tableau 23 : Effet de l'emploi saisonnier et à temps partiel (Allier) | 115 |
| Tableau 24 : Effet de l'emploi saisonnier et à temps partiel (Cantal) | 115 |
| Tableau 25 : Effet de l'emploi saisonnier et à temps partiel (Haute-Loire) | 116 |
| Tableau 26 : Effet de l'emploi saisonnier et à temps partiel (Puy-de-Dôme) | 116 |
| Tableau 27 : Récapitulatifs des deux corrections | 116 |
| Tableau 28 : Ratios de branche corrigés (euros par emploi, 2017, Allier) | 117 |
| Tableau 29 : Ratios de branche corrigés pondérés (Allier) | 117 |
| Tableau 30 : Ratios de branche corrigés (euros par emploi, 2017, Cantal) | 118 |
| Tableau 31 : Ratios de branche corrigés pondérés (Cantal) | 118 |
| Tableau 32 : Ratios de branche corrigés (euros par emploi, 2017, Haute-Loire) | 119 |
| Tableau 33 : Ratios de branche corrigés pondérés (Haute-Loire) | 119 |
| Tableau 34 : Ratios de branche corrigés (euros par emploi, 2017, Puy-de-Dôme) | 120 |
| Tableau 35 : Ratios de branche corrigés pondérés (Puy-de-Dôme) | 120 |
| Tableau 36 : Effets directs de la dépense touristique : coefficients d'emploi, de valeur ajoutée et de masse salariale | 121 |
| Tableau 37 : Effets directs des dépenses des visiteurs de la CPFL sur l'économie des départements | 122 |
| Tableau 38 : Répartition de l'emploi entre sphères productive et présenteielle en 2017 | 126 |
| Tableau 39 : Données pour le calcul des coefficients d'induction dans les départements | 127 |
| Tableau 40 : Effets directs et induits des dépenses des visiteurs de la CPFL sur l'emploi dans les départements | 128 |

III Liste des figures

| | |
|--|-----|
| Figure 1 : Détermination des effets directs, indirects et induits..... | 31 |
| Figure 2 : Chaîne des Puy-Faille de Limagne | 47 |
| Figure 3 : Sources des informations utilisées pour estimer les effets locaux de la visite de la CPFL | 97 |
| Figure 4 : Corrélation entre emplois productifs et emplois présentsiels..... | 124 |
| Figure 5 : Corrélation entre population résidente et emplois dans la sphère présenteielle | 124 |

A compléter

I. Introduction

A. Le contexte

Le Conseil départemental du Puy-de-Dôme (CD63) mène en relation étroite avec divers partenaires (PNR des Volcans d'Auvergne, services de l'Etat, Fondation privée, ...) une politique de gestion et valorisation des sites emboîtés du puy de Dôme et de la Chaîne des Puys - Faille de Limagne (CPLF).

Cette politique implique des investissements matériels et immatériels, des dépenses de fonctionnement à travers des plans de gestion, et s'accompagne de labellisations ou autres distinctions.

Se sont ainsi succédées : la labellisation en 2008 du site du puy de Dôme en tant que « Grand site de France» (politique assortie d'un plan d'aménagement et de gestion, et incluant, avant le renouvellement du label en 2014, des travaux lourds tels que l'aménagement du sommet du puy de Dôme et la réalisation du train à crémaillère), et l'inscription en 2018 du site de la CPFL sur la liste UNESCO du Patrimoine mondial (assortie elle aussi d'un plan de gestion).

Le travail de fond sur l'ensemble du site, repris dans les divers plans de gestion, comprend de multiples actions sur des thématiques telles que la protection des milieux, la gestion des paysages ou la gestion des accès et des circulations. Un travail de communication est mené en direction des résidents du département et en direction de tous ceux qui seront sensibles à l'attractivité des sites (nouveaux actifs, touristes, etc.).

L'ensemble de ces actions sera appelé la « politique de référence ».

Les acteurs à l'origine de cette politique de référence souhaitent évaluer au mieux les retombées économiques de cette politique publique sur le territoire, et disposer de méthodes permettant d'actualiser ultérieurement cette connaissance.

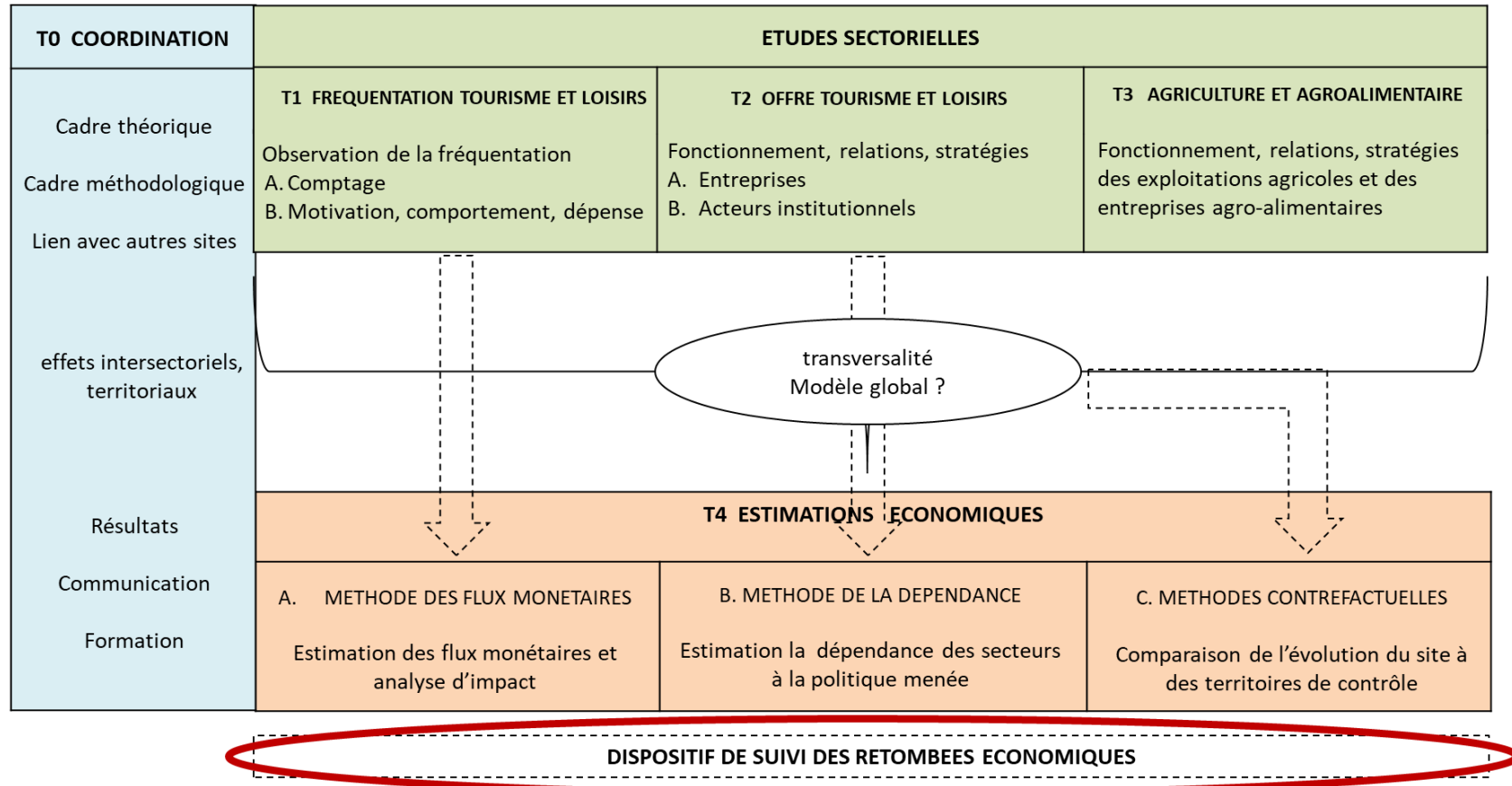
Des chercheurs et enseignants-chercheurs de l'UMR Territoires se sont proposé de mener une recherche qui contribue à répondre à ces types d'attentes. Recherche-action, le projet est mené en relation avec les acteurs de la politique menée et avec des acteurs des principaux secteurs concernés.

Parmi les effets attendus de la politique, certains sont sectoriels (tourisme, agriculture, etc.), d'autres plus globaux (image, attractivité, dynamique de développement) ayant également des retombées sectorielles. Certains effets seront plus facilement quantifiables que d'autres. Enfin ils porteront sur diverses échelles, de celle du site à celle de l'Auvergne, voire au-delà.

Les objectifs initiaux de Valo-Sites se répartissent ainsi :

- Observer, décrire, qualifier et mesurer des évolutions socio-économiques, possibles effets des politiques d'aménagement de gestion et de valorisation d'un site naturel protégé (le site CPFL).
- Tester, choisir, appliquer et décrire les méthodologies employables. Les questions méthodologiques se présentent à divers niveaux : quelle méthode praticable et reproductible pour connaître la fréquentation totale d'un site aussi étendu ? quelles méthodes de mesure d'effets économiques directs, voire indirects et induits ? Quelle description d'effets généraux (attractivité, développement) ?
- Analyser puis décrire les moyens et l'organisation à mettre en œuvre pour assurer ce travail et pour le reproduire ou l'actualiser ultérieurement dans un laps de temps donné pour le cas de la CPFL (faire ou faire faire.).
- Ces travaux doivent également permettre de mettre en avant les conditions de transposition des différentes méthodes mobilisés à d'autres sites comparables.

B. L'organisation générale des tâches du projet



C. Les objectifs des tâches T1-B et T4-1

L'objectif principal de la recherche, concernant ces tâches, est de mesurer des indicateurs de l'impact économique de la fréquentation du site à des fins de tourisme et de loisirs, afin de pouvoir ultérieurement (au-delà du projet Valo-sites) en suivre l'évolution dans la durée, puis de tenter de comprendre dans cette évolution la part des politiques de gestion et de valorisation.

L'objectif n'est pas de mobiliser des méthodes pour *interpréter* l'évolution future des indicateurs observés. Cependant nous faisons l'hypothèse que la recherche de la meilleure connaissance possible de la fréquentation permet une meilleure compréhension des *mécanismes* d'impact, et donc *des raisons de leur évolution*.

La tâche T1-B a pour objectif la connaissance de la fréquentation, et la tâche T4-1 a pour objectif la mesure d'un impact économique à *partir de données collectées par T1-B*.

Concernant T1-B, la collecte d'informations sur la fréquentation (nécessaires à la mesure de l'impact économique) permet également d'éclairer une politique d'offre et d'aménagement. Ce sera un objectif secondaire de la tâche.

II. Problématiques, cadres théoriques et méthodologiques

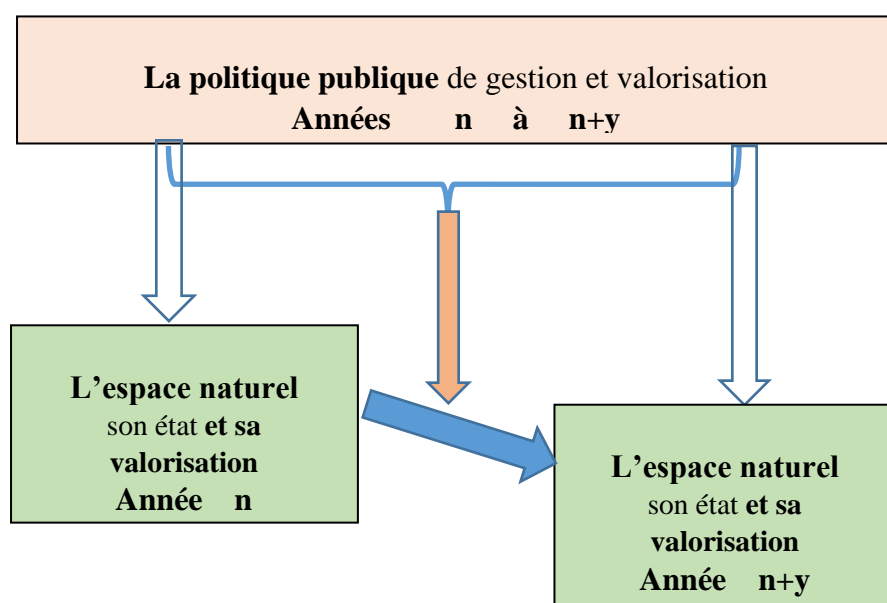
A. Problématique générale : valeurs d'un espace naturel, politiques de gestion et valorisation, évaluations diverses

On adoptera le terme de « gestion » d'un espace dans le sens le plus large du concept, incluant les dimensions d'aménagement, de gestion courante, de promotion, de gouvernance. On considèrera que d'autres termes tels que ceux de « protection » ou de « valorisation » décrivent plutôt des finalités de cette gestion.

1. Distinguer état (valeurs) de l'espace, évolution de cet état, et évaluation de politiques menées

Ce sont les politiques publiques et les actions, publiques ou privées, de gestion (passées présentes et futures) qui posent la question de la valeur de ces espaces et de l'évaluation de ces politiques et actions. Chacun de ces concepts, valeur et évaluation, a lui-même plusieurs acceptions (CREDOC 2008, CGDD 2010, Binet, de Battisti et al. 2011, Ministère_environnement 2017, Duval, Smith et al. 2019).

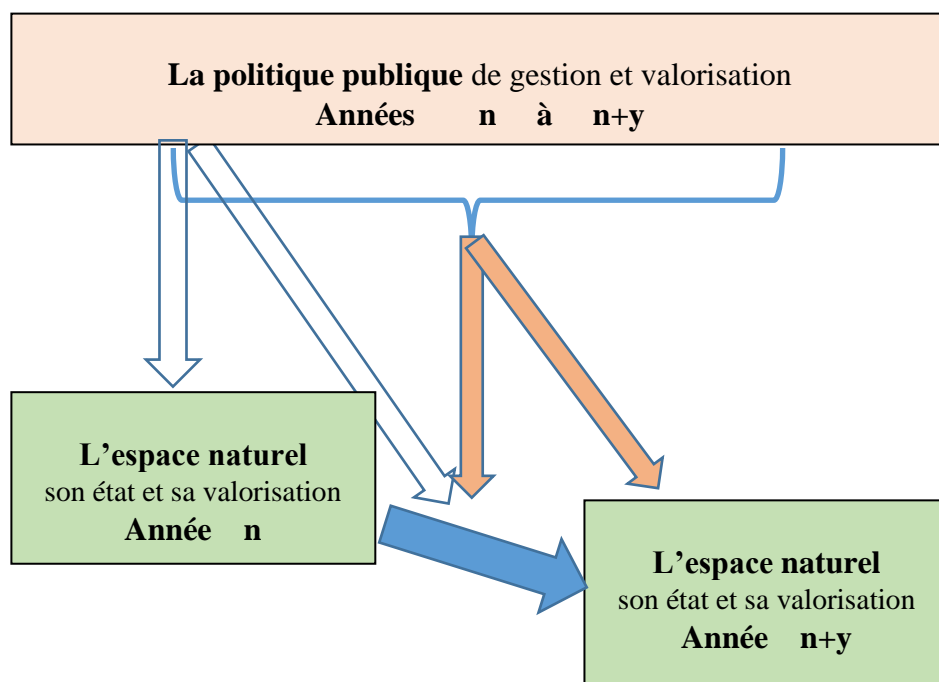
Une distinction essentielle, et pas toujours claire dans les références parcourues, est à faire entre valeurs de l'espace, variations de cette valeur et effets d'une politique menée sur cet espace. Le schéma suivant illustre ces articulations :



Le premier raccourci erroné serait d'attribuer à la politique en cours sur l'espace considéré une valeur de cet espace constatée/mesurée à un instant t (année n) : le premier correctif est de ne considérer que *l'évolution* de cette valeur entre deux moments choisis (par exemple un moment choisi avant la première application de la politique et un moment actuel).

En fait la politique menée à l'année n agit instantanément sur les *états* de l'espace et sur les conditions de sa valorisation, ce faisant elle agit aussi sur les *conditions d'évolution* de ces deux composantes (c'est le sens des deux flèches blanches).

Ainsi l'état final de l'espace et de sa valorisation sera la résultante de toutes ces inflexions années après années (flèches orange), sans oublier qu'il est aussi la résultante de causes diverses qui ne tiennent pas à la politique examinée, ce sera le point suivant.

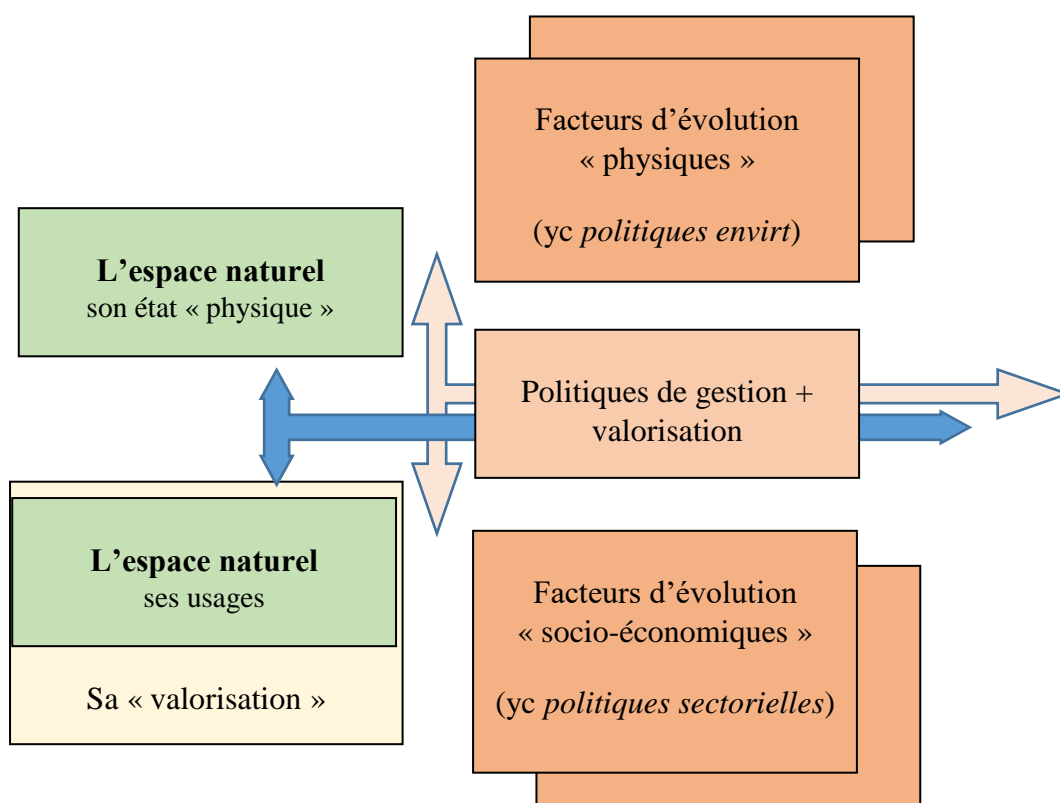


La deuxième approximation serait d'attribuer à la politique menée une causalité totale des variations de valeur constatée. C'est le problème de l'évaluation des « *effets propres* » d'une politique c'est-à-dire faisant la part des autres causes d'évolution possibles : ce point fait l'objet de la section qui suit.

2. Distinguer les effets propres, distinguer gestion et valorisation

La question de savoir comment distinguer les « *effets propres* » d'une politique dépasse le champ de Valo-sites.

Néanmoins pour mieux cadrer les travaux il est utile de dissocier même de façon imparfaite les dimensions du sujet pour mieux comprendre à quelle *diversité de causes potentielles* l'évolution du site et de ses valorisations peut être soumise.



3. Valeurs d'un espace naturel

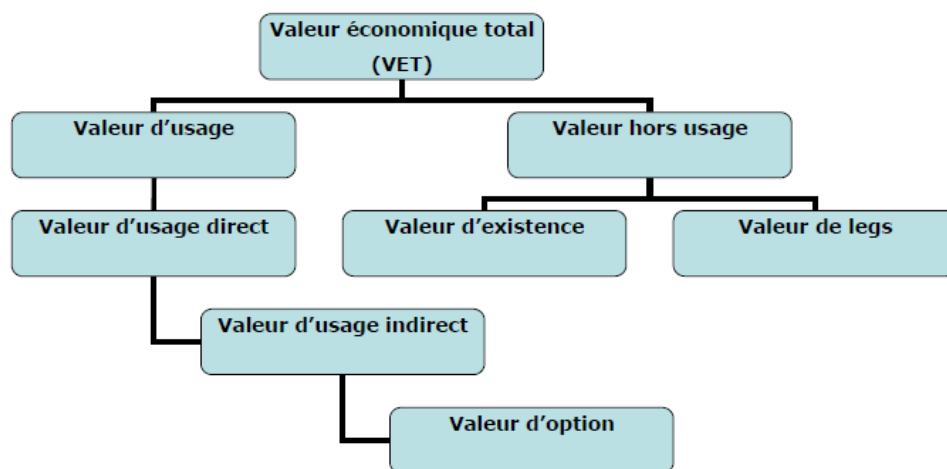
Il est difficile de dissocier la définition d'une notion telle que celle de « valeur » et le questionnement que l'on poursuit. C'est ainsi que CREDOC, Dujin et al. (2008) exposent successivement 4 façons d'aborder la question de la valeur économique d'un espace naturel :

- Analyser la contribution d'un écosystème à l'activité économique
- Evaluer les retombées économiques de la protection de l'espace
- Déterminer la « valeur économique totale » d'un bien public
- Etablir le bilan coûts-bénéfices de la protection d'un espace

Seules les approches 1 et 3 donnent une définition hors considérations sur la politique envisagée, et l'approche 3 est la typologie la plus complète des différentes valeurs ... économiques.

La valeur économique et sociale des espaces naturels protégés

La valeur économique totale, d'après S. Pagiola, K. Von Ritter, Joshua Bishop [2004],
Assessing the Economic Value of Ecosystem Conservation



CREDOC, 2008

NB : cette typologie est la plus complète : par exemple les services économiques écosystémiques y sont présents, portés au titre de la « valeur d'usage indirect ».

Mais une approche non-économique aura du mal à accepter de ranger une « valeur symbolique » par exemple dans la notion de « valeur hors usage » ... du simple fait que la typologie exposée s'inscrit dans un paradigme de monétarisation (notamment pour aider à une prise de décision sur base mathématique) que d'autres points de vue peuvent vouloir contourner.

Le guide proposé par Ministère_environnement (2017) s’attache à éclairer les acteurs des « territoires d’exception ». Sous ce vocable il vise explicitement les territoires objets de deux politiques : celle des grands sites de France (OGS) et celle de l’inscription à la liste UNESCO du Patrimoine mondial.

Ce guide propose de « construire un système d’évaluation utile aux gestionnaires et acteurs des territoires d’exception. Cette évaluation repose les trois valeurs du développement durable : valeur économique, valeur sociale et valeur environnementale, ainsi qu’une valeur territoriale qui traduit les dynamiques territoriales générées ».

Gamme des valeurs étudiées dans ce guide

| La valeur économique | La valeur sociale | La valeur environnementale | La valeur territoriale |
|------------------------------------|--|---|--|
| Retombées économiques | Effets sur la qualité de vie et les impacts sociaux (bien-être, accessibilité, mobilité) | Protection de l’environnement (paysage et biodiversité) | Valorisation des représentations du territoire et des usages |
| Création de richesses et d’emplois | Cohésion sociale | Préservation des « équilibres environnementaux » | Qualité et dynamiques de la gouvernance |

B. La valorisation économique des espaces protégés

De façon croissante, mais pas partout, les gestionnaires d'espaces protégés associent un objectif de valorisation économique à l'objectif initial de protection et de gestion de ces espaces (Charlot, O., 2007 ; Fredman, P. and M. Yuan, 2011), ou du moins espèrent-ils que la protection stimulera le développement (Duvivier, 2021).

Une question est donc de savoir si les politiques de protection et de bonne gestion, et particulièrement la « labellisation » de ces politiques accroissent les retombées économiques (Hall, C. M. and R. Piggin, 2001 ; Prud'homme, R., et al., 2008 ; Tisdell, C. and C. Wilson, 2002 ; Talandier, M., et al., 2020). Une revue de la littérature à ce sujet, menée dans le cadre du projet Valo-Sites aboutit à la conclusion qu'à l'échelle mondiale, l'effet des sites naturels protégés sur le développement n'est pas confirmé (Duvivier, 2021).

Les raisons tiennent à une grande diversité des situations, sur le plan des caractéristiques intrinsèques des sites, de leur histoire, des contextes territoriaux, et des objectifs mêmes des politiques qui sont menées localement.

En effet ces politiques ne s'accompagnent pas toutes d'une volonté de développement, et même les distinctions obtenues ne sont pas toujours promues par les acteurs locaux. Gravari-Barbas (2002) estime que « *le patrimoine, [...] peut avoir un impact positif sur les territoires, à condition que cette ressource fasse l'objet d'une construction territoriale* ».

Cette condition est loin d'être généralement remplie, selon une étude menée par Rebanks Consulting Ltd. and Trends Business Research Ltd (2009) sur les politiques affichées par les sites inscrits à la liste UNESCO du patrimoine mondiale, montrant que 600 sur les 878 sites analysés *n'ont pas d'objectif de valorisation*. Les auteurs affirment ainsi que « *les impacts [économiques] de l'inscription [Unesco] sont rarement accidentels ou inintentionnels* ».

Dans la même perspective, Talandier et al. (2020) classent les territoires disposant de sites patrimoniaux exceptionnels en trois types, dont l'un où le site n'est qu'une ressource latente ; au-delà des facteurs structurels, ils concluent à l'importance de la capacité des acteurs à se coordonner pour valoriser cette ressource.

C. La part importante de la fréquentation de tourisme et loisirs dans la valorisation économique (et sociale) des espaces naturels

Le tourisme est l'un des facteurs majeurs d'un développement potentiel autour des espaces naturels protégés (Lazzarotti, O., 2000 ; Mayer, M., et al., 2010 ; Stynes, D., 2005 ; Charlot, O., 2007 ; Puhakka, . and J. Saarinen, 2013 ; Cassel, S. H. and A. Pashkevich, 2014). Les touristes peuvent porter attention aux statuts d'espaces protégés, qui influencent alors leurs décisions de visite ou de choix de destination (Deng, J., et al. ; 2002 ; Reinius, S. W. and P. Fredman, 2007).

Diverses questions, telle que la cohabitation difficile entre fréquentation touristique et autres usages des espaces, ou bien la diminution des moyens publics disponibles, suscitent une demande croissante d'évaluation des apports économiques des usages récréatifs et touristiques (Stynes, D., 2005). Les progressions envisageables dans ce domaine au niveau des effets économiques directs, sont un accroissement du nombre de visiteurs, et/ou de leurs durées de séjour, et enfin de leurs dépenses (ce qui peut tenir également à une modification de l'offre et de ses prix, ou du profil des visiteurs).

Duvivier (2021) relève les diverses causes expliquant l'effet ou le manque d'effet dans le secteur du tourisme, des politiques menées globalement :

- effort ou non de promotion,
- présence ou non d'équipements et de services touristiques,
- force relative d'attraction du statut d'espace protégé par rapport aux qualités intrinsèques du site et à sa notoriété antérieure (par exemple la politique des Grand Sites de France vise -par construction- des sites dont la fréquentation touristique est déjà particulièrement importante avant le processus de labellisation), etc.

D. Les dimensions sociales et économiques de la fréquentation d'un site naturel (protégé)

1. Enjeux

Les espaces naturels font l'objet, de façon diverses et inégales, de deux types de préoccupations :

- 1- Leur valorisation par des activités économiques et/ou résidentielles, parmi lesquelles se situent les activités de tourisme et de loisirs. Le sens du terme « valorisation » est mal défini, il dépend en particulier des cadres théoriques qui tentent -ou non- de donner une valeur à des usages non marchands de ces espaces¹. Dans la tâche 4-1 du projet Valo-sites on ne s'attache pas aux valorisations non-marchandes.
- 2- Leur entretien et leur protection le cas échéant, ce à quoi certains acteurs restreignent le terme de « gestion » des espaces naturels. *Pour les spécialistes du champ de la gestion, il serait approprié de dire « gestion du site, comprenant aménagement, entretien, protection et valorisation ».*
On parle d'espaces naturels protégés lorsque la protection fait l'objet de mesures réglementaires.

Pour sa part *l'aménagement* d'un site (comme sa gestion au sens large) peut avoir pour objectif l'une ou l'autre des finalités (valorisation ou gestion/protection) ou les deux.

Le projet s'attache à un ensemble de politiques « d'aménagement, de gestion et de valorisation » d'espaces naturels protégés, dont le site de la Chaîne des Puys-Faille de Limagne est le cas support.

Enjeux de la valorisation

¹ La littérature scientifique s'est attachée notamment aux concepts « d'aménités », puis de « services éco-systémiques ».

Les politiques de gestion occasionnent des coûts pour le gestionnaire et la mise en place de protection occasionne *des coûts* pour les acteurs locaux et *des résistances*

La valorisation est alors :

- Un facteur potentiel *d'engagement*, spontané ou suscité, des acteurs locaux concernés
- Un facteur de *compensation des coûts subis*, à l'échelle locale comme à l'échelle des gestionnaires du site global
- Un moyen potentiel de faire participer *directement* les visiteurs aux coûts d'entretien du site (AFIT 2000)

2. Dimensions marchandes et non marchandes de la valorisation

Une activité telle que le tourisme (de même que celle de loisirs) possède(nt) des composantes marchandes et non marchandes :

- Vues du côté du touriste, une part des activités de celui-ci et de sa satisfaction proviennent de composantes non marchandes (contemplation, promenade, etc.).
- Vues du côté de l'offre touristique, une part de celle-ci est composée d'infrastructures, voire de services, publics ou privés, gratuits (chemins, espaces verts, animation, etc.) soit en accès direct pour le touriste, soit utilisés par des entreprises touristiques : ces infrastructures et services participent donc à la « valeur » du produit touristique sans que leur apport ne soit directement répercuté dans un prix (Petiot 2004). Un éventuel différentiel de prix pourrait en rendre compte partiellement.
- Enfin les activités du touriste comme celles des entreprises touristiques, comme celles des fournisseurs d'infrastructures et de services d'accompagnement, créent des *externalités* positives ou négatives qui ne sont par définition pas comptabilisées dans les échanges marchands.

3. La relation du touriste/excursionniste à l'espace naturel : directe ou via le « triangle des services »

L'économie des services (Gadrey 1994) ainsi que le management de la production de service (Eiglier and Langeard 1987) ont mis en évidence le caractère triangulaire des relations de service qui se nouent entre un bénéficiaire, un prestataire de service et aussi ce qui est appelé le « *bien support du service* » : ce dernier peut être un objet (l'automobile que le garagiste répare), une infrastructure matérielle (la chambre d'hôtel), ou autre.

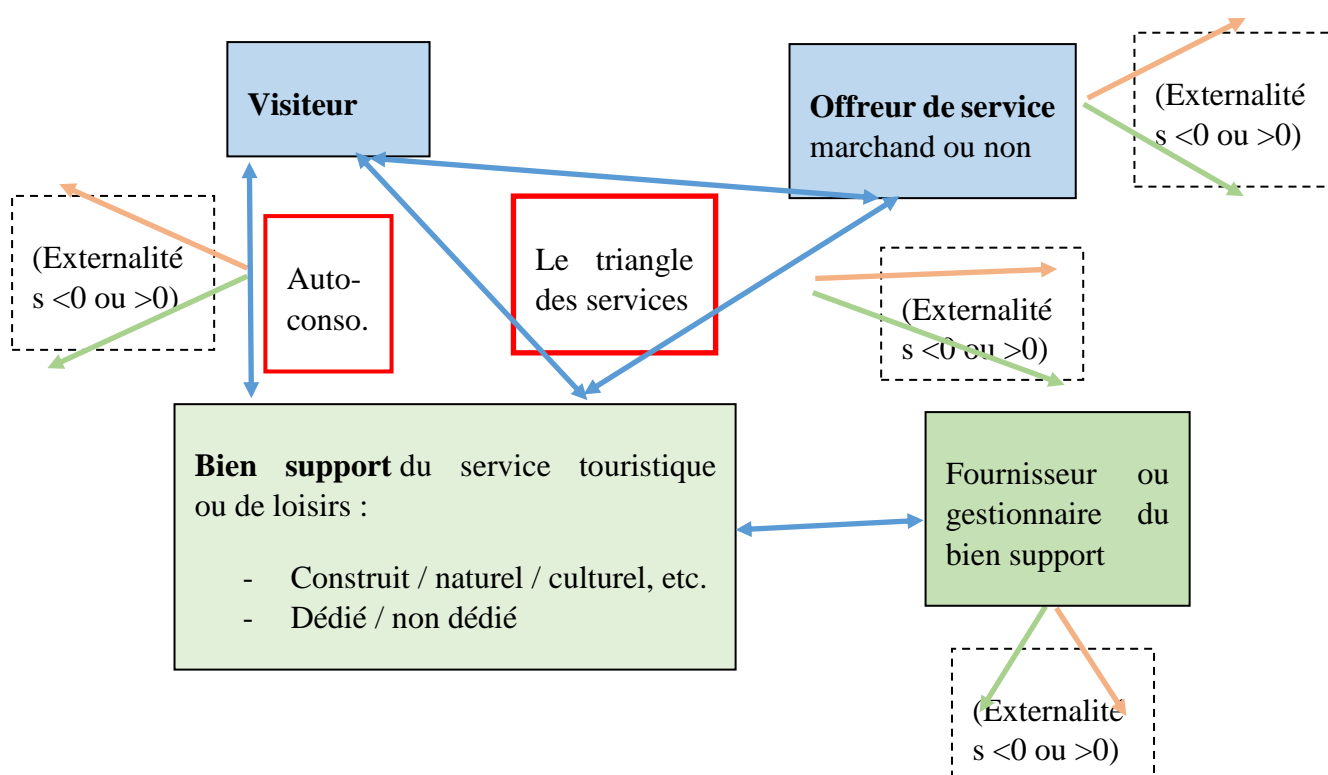
Ainsi l'espace naturel ou un attribut de celui-ci peuvent être les « biens-supports » d'un service touristique ou de loisirs.

Ce service peut être autoproduit et autoconsommé lorsque le bénéficiaire (ici le visiteur) profite directement du ruisseau, de la qualité de l'air, du paysage. Lorsqu'il profite d'un sentier on identifie déjà l'aménageur, ou celui qui entretient, qui balise le sentier comme un *fournisseur* de bien-support.

L'accompagnateur en moyenne montagne sera un *prestataire de service* qui mobilise le bien-support-sentier.

Dans le schéma ci-dessous sont figurées les deux types de relations de service : l'une en autoconsommation, l'autre sous forme du triangle des services. Le schéma fait figurer l'éventuel offreur (prestataire) de service et l'éventuel fournisseur (ou gestionnaire) du bien support.

En référence à la section précédente, il signale que les relations peuvent être *marchandes ou non* et il fait figurer également les éventuelles *externalités* économiques attachées à chacun des acteurs ou aux relations d'autoconsommation ou de service.



Ce modèle se décline à chaque relation de service. Ainsi un visiteur-touriste fera l'objet de relations d'hébergement, de restauration, de loisirs divers avec *pour chacune de ces relations de service, des prestataires et des biens-supports particuliers*.

Mais il a été proposé une généralisation analogique consistant à considérer une « destination » touristique comme bien-support agrégé, et l'ensemble des acteurs de cette destination comme un quasi-prestataire unique (Meyronin and Valla 2004, Marsat, Brault et al. 2009, Marsat, Bonniot et al. 2013).

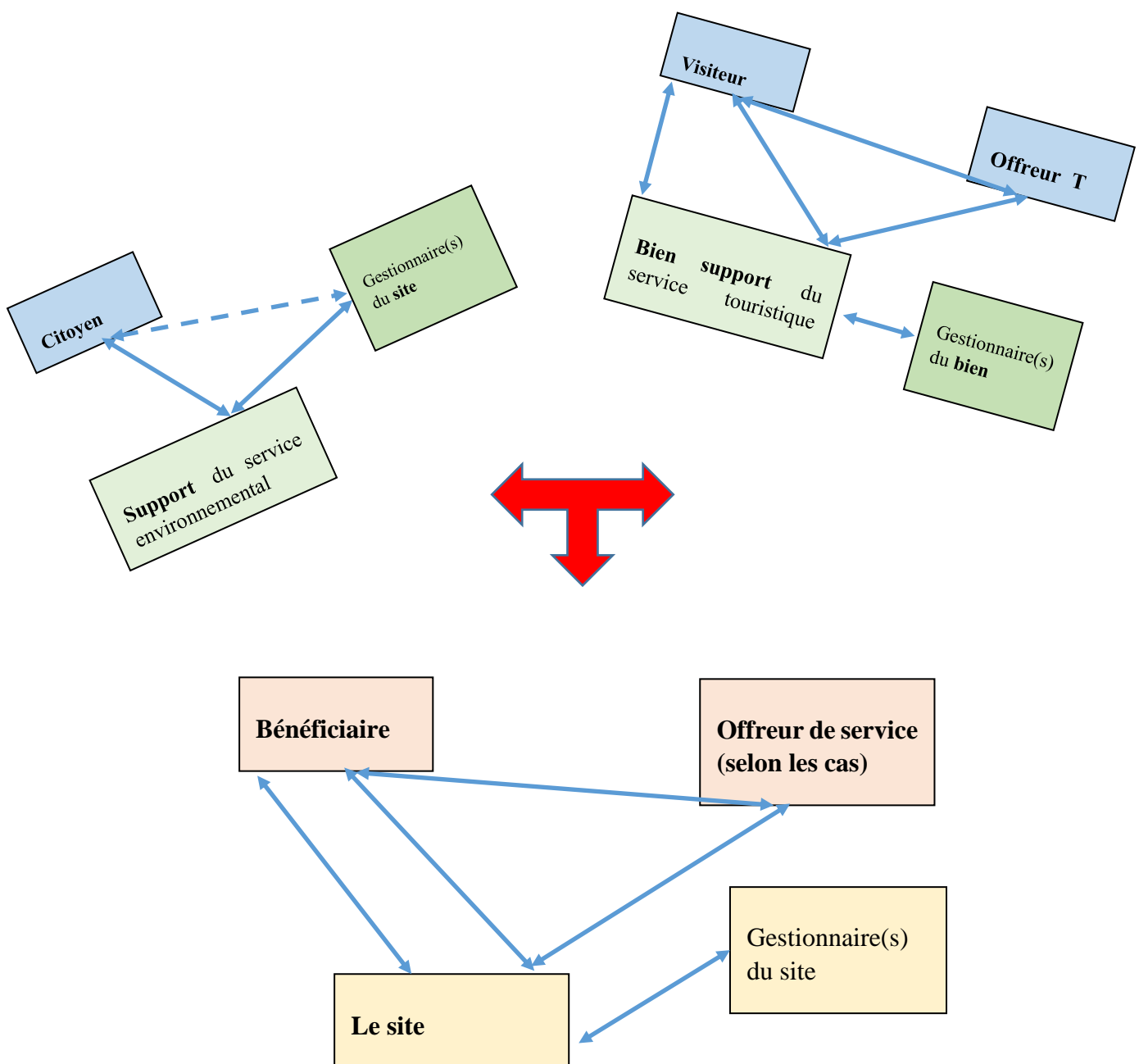
Un espace naturel tel que la CPFL constitue alors un bien-support qu'on dira « intermédiaire » : En effet il n'y a pas superposition entre l'espace naturel cité et la ou les « destinations touristiques » vues par des touristes qui logeront éventuellement en dehors de l'espace CPFL et qui visiteront d'autres espaces. C'est l'objet du point II.4.4.

Auparavant une généralisation de la référence aux services (et au concept de triangle des services), est également éclairante. Il s'agit de dépasser le champ de la valorisation par la fréquentation de tourisme et de loisirs.

En effet l'activité de gestion du site représente également des fournitures de services à des bénéficiaires divers.

La gestion à des fins de protection est notamment, entre autres, un service rendu aux bénéficiaires intéressés à toute valeur de non-usage (valeur de legs, valeur d'existence). C'est plus largement une approche à relier au champ des « services écosystémiques ».

La gestion à des fins de valorisation pédagogique est un autre service rendu, en partie à d'autres bénéficiaires. Etc.



4. Géographie du phénomène de tourisme et loisirs liés à la visite d'un site

Deux notions sont mobilisées ici : celle de *destination touristique* et celle d'*attraction touristique*.

a) La « destination touristique »

Les diverses définitions de la destination touristique forment un continuum qui va du sens premier du terme (est « destination » ce vers quoi on va) jusqu'à une liste des attributs de toute destination, considérés comme essentiels. Ainsi :

- la définition minimale serait : *est destination touristique « l'aire géographique » que le touriste a identifiée et choisie comme lieu de séjour (fixe ou itinérant) et de pratiques.*

- la définition étendue serait : *est destination touristique l'ensemble d'attributs et de services concernés par / contribuant au séjour du touriste.*

Par exemple voici deux définitions parmi celles qu'a rassemblées Phetvaroon (2006) :

- Minimale : « *A place where people spend their holiday* » (Laws, 1995) mais celui-ci ajoute déjà une courte liste : « its elements entail place, people, and holiday ».
- Maximale : « *An amalgam of individual products and experience opportunities that combine to form a total experience of the area visited.* » (Murphy, Pritchard, & Smith, 2000)

Ainsi la destination est souvent appréhendée comme un système qui apporte une réponse aux besoins des touristes (Botti, Peypoch, Solonandrasana, 2008). La destination touristique peut être décrite comme une aire géographique qui est en mesure d'offrir un produit touristique, c'est-à-dire un ensemble de services supports qui gravite autour d'activités ou expériences inhabituelles pour le touriste. (Framke, 2002).

Il est clair que les visiteurs d'un site peuvent avoir choisi une destination plus vaste que le site, ne serait-ce que pour se loger.

De plus s'ils viennent dans cette destination pour visiter plusieurs sites, s'ils ne viennent sur le site considéré que pendant une partie de leur séjour, alors le site a un autre statut : celui d'« *attraction touristique* ». Cela sera l'objet de la section suivante.

Il n'y a pas de destination touristique unique à définir en lien avec la fréquentation du site. Selon ses objectifs, la part du site parmi ceux-ci, chaque touriste définit sa destination et le poids du site dans cette définition sera variable.

Il sera possible de décrire la géographie de la fréquentation au niveau des pratiques d'hébergement, restauration, visites : une *enveloppe maximale*, notamment pour les lieux d'hébergement, apparaîtra (car la distance au site joue dans l'établissement d'un trajet maximum). Il est possible de définir la notion « *d'enveloppe maximale des destinations individuelles liées au site* ». Il sera possible aussi de décrire des gradients dans cette géographie.

b) « L'attraction touristique »

Une attraction touristique est l'un des éléments qui « attire » le touriste (de son domicile habituel) vers sa destination, c'est "un élément [...] qui tire le touriste hors de son environnement naturel" (Lew, 1987).

Comme au sujet de la « destination » les définitions de l'attraction touristique vont de la définition minimale qui prend en compte les éléments formels attractifs (le climat, la beauté naturelle, un monument réputé, un contexte favorable pour un sport donné) à des définitions qui incluent des éléments organisationnels sans oublier la dimension cognitive par laquelle le touriste prend en compte cette attractivité (Leiper 1990).

Cette notion d'attraction a une portée qui n'est pas que géographique. Elle a des implications en marketing notamment. Leiper classe les attractions selon qu'elles sont *primaires* et motivent alors le choix de la destination ; *secondaires*, elles sont soupçonnées, mais relativement mal connues ou moins appréciées, et ne déclenchent pas à elles-seules le voyage, enfin, *tertiaires*, elles sont ignorées du touriste avant le départ mais contribuent sur place à sa satisfaction.

Il n'y a pas d'universalité de ce classement qui dépend des sensibilités de chacun, une attraction primaire pour l'un pouvant devenir secondaire pour l'autre (Bonniot, Bouchaud et al. 2010). Cette typologie rejoint l'opération qui sera effectuée plus tard dans la présente tâche du projet et qui consiste à estimer la force de « l'effet site » dans le choix de destination de chaque touriste rencontré.

Enfin la dimension cognitive signalée par les auteurs, permet de ne pas oublier le rôle de la notoriété, de l'image et des distinctions acquises par le site, dans le système touristique¹.

c) Destination(s), attraction(s) et territoire(s) d'impact : imbrications

Comme cela a été exposé ci-dessus, même dans le cas où un site est une attraction primaire pour un visiteur, ce dernier définira sa « destination » possiblement sur une aire plus vaste incluant les fonctions d'hébergement, et peut-être d'autres attractions (secondaires). Un site tel que la CPFL peut être une destination à part entière puisqu'on peut y être hébergé et nourri, mais selon les visiteurs considérés, le site peut être une destination à part entière ou une attraction au sein d'une destination plus vaste.

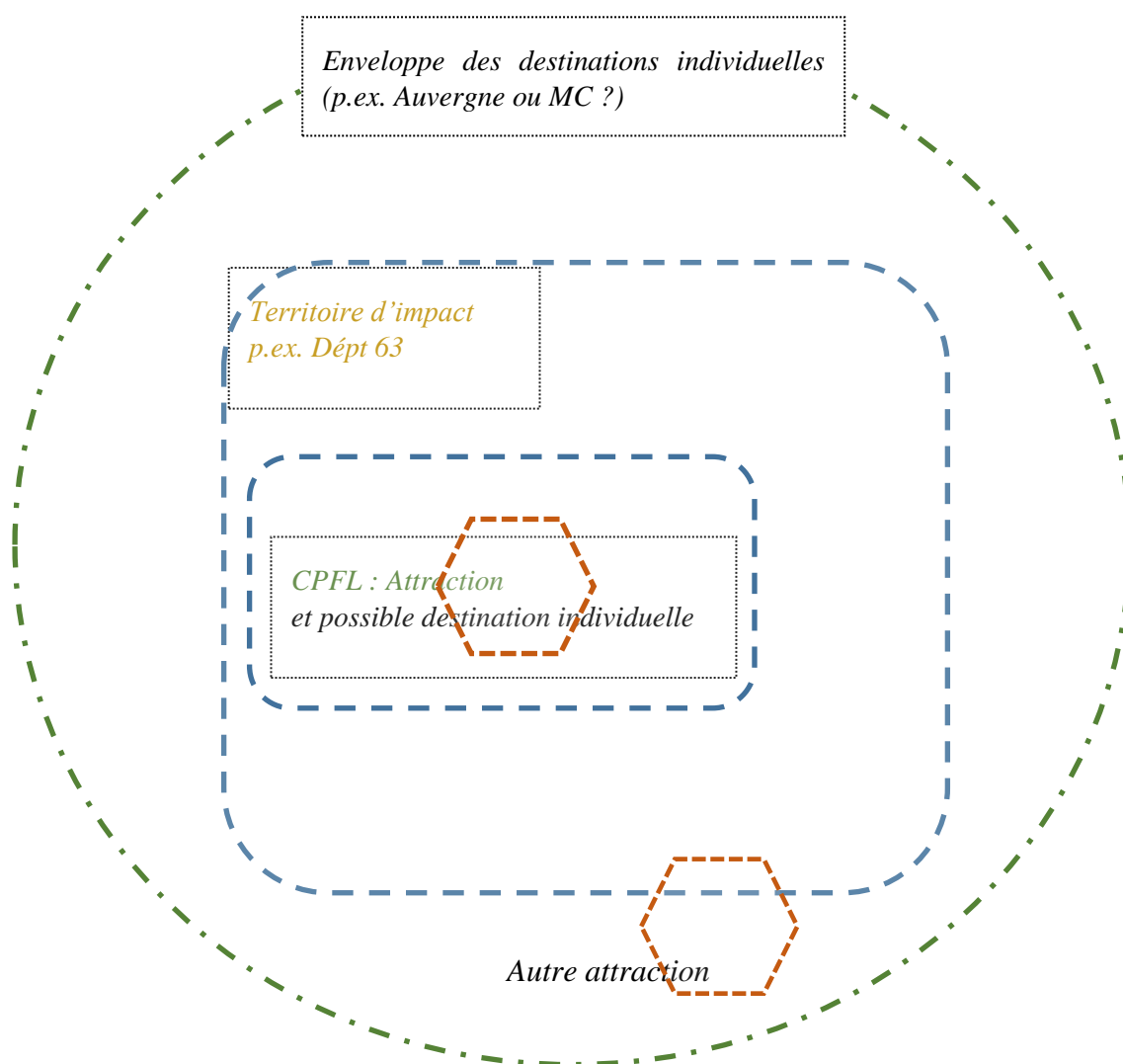
De plus il peut exister d'autres attractions au sein de la même « enveloppe des destinations individuelles liées au site » : ces autres attractions peuvent susciter une partie de dépenses des visiteurs du site étudié (la CPFL).

¹ Ainsi le schéma proposé ci-dessus en VII.4.3 pourrait inclure un lien (une flèche) allant de la CPFL à un triangle de service donné, pour les cas où la réputation de la CPFL intervient même si physiquement ce triangle donné ne met pas en jeu l'espace CPFL

Pour aller plus loin, mais cela ne concerne que les préoccupations de gestion d'offre, de management de destination touristique, il y a des superpositions géographiques entre cette « *enveloppe des destinations* » liées au site étudié (ici la CPFL) et les « *enveloppes de destinations* » liées à d'autres attractions (par exemple le Massif du Sancy), l'ensemble étant réparti selon les « effets » de telle ou telle attraction dans les choix de destination, puisque une attraction (comme la CPFL) peut être primaire pour tel touriste, secondaire ou tertiaire pour tel autre.

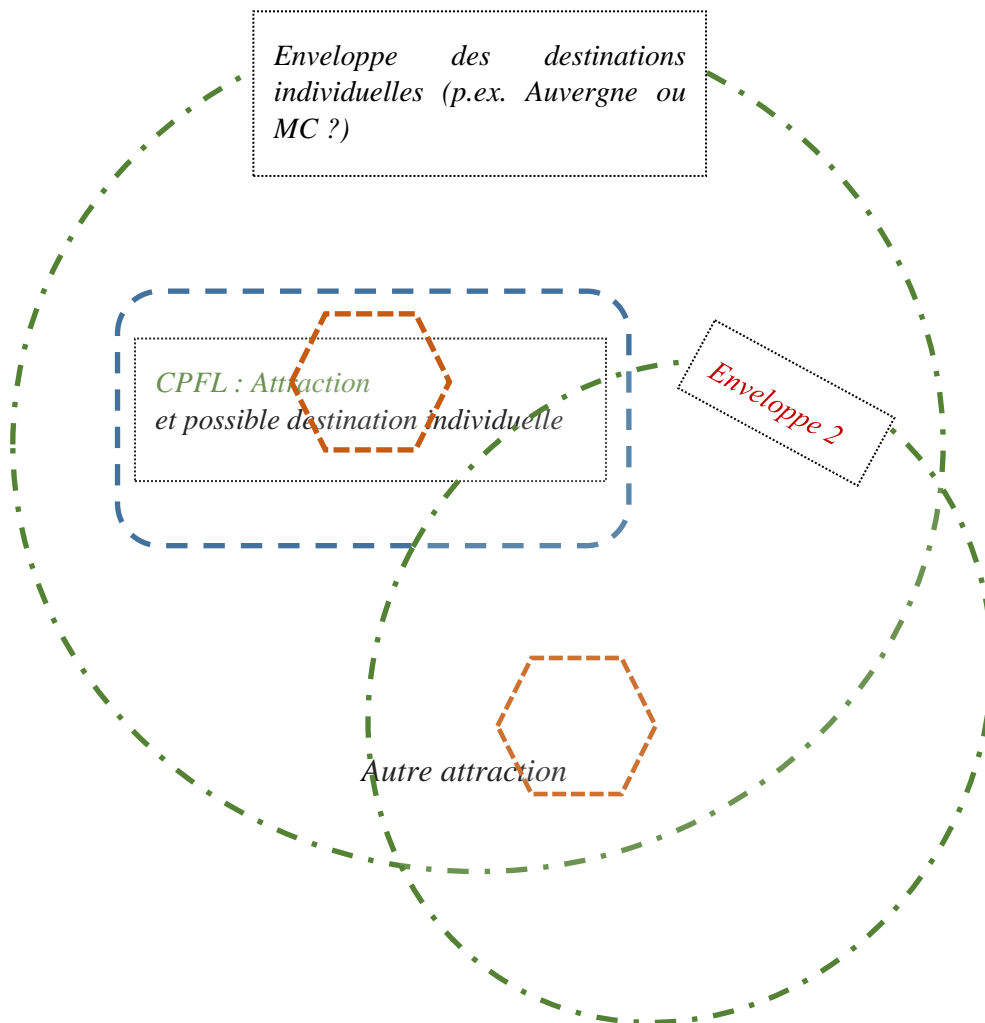
Enfin la notion de *territoire d'impact* est celle du territoire librement choisi pour évaluer les retombées du phénomène étudié sur ce territoire ¹. On notera qu'un « territoire d'impact » peut être choisi qui n'inclut pas géographiquement le site (la CPFL).

Ainsi les aires à considérer sont-elles représentables ainsi :



¹ On verra que dans le cas de la CPFL ce territoire d'évaluation peut être celui des intercommunalités sécantes avec la CPFL, ou bien le département 63 ou l'Auvergne, etc.

Les autres attractions peuvent être « primaires » pour certains visiteurs du site (CPFL), on peut donc considérer « l'enveloppe des destinations » correspondante : c'est une deuxième enveloppe qui chevauche la première. On en trouvera des exemples (Massif du Sancy).



E. La connaissance de la fréquentation des espaces naturels (protégés) : objectifs, dispositifs d'observation, définitions et notions-clés

1. Objectifs

La connaissance de la fréquentation des espaces naturels peut répondre à divers objectifs éventuellement de façon conjointe.

C'est une question qui est habituellement posée pour le compte des acteurs en charge de la gestion des espaces et/ou de la valorisation de ces espaces, ou de politiques plus larges (rarement pour le compte des visiteurs eux-mêmes).

On peut donc distinguer parmi les objectifs :

- ceux qui s'attachent à mieux protéger l'espace considéré, à gérer la diversité des usages : connaître les pratiques des visiteurs et leurs conséquences sur l'état du milieu, etc. calculer la « capacité de charge » du milieu. Eventuellement connaître les motivations, les représentations, les demandes des visiteurs pour rapporter ces informations à la protection de l'espace, à la gestion des usages.
- ceux qui s'attachent à la valorisation de l'espace, ce qui peut se résumer à une sorte de fonction de type « marketing » prise au sens complet (c'est à dire: connaître la demande (marchande ou non) et adapter l'offre) et appliquée à toutes les dimensions possibles de la notion de « valorisation » (sociale, économique). Il s'agit alors de priorités différentes : les motivations, les déterminants du choix de destination, les satisfactions et les attentes sont à connaître pour attirer et/ou satisfaire. L'ensemble des connaissances, y compris sur les pratiques, les activités, les infrastructures et les moyens employés par les visiteurs sont utilisées également pour éclairer des politiques d'offre.
- Ceux qui s'attachent à l'évaluation des politiques menées sur l'espace considéré, ce groupe d'objectifs pouvant « boucler » avec les objectifs précédents selon qu'on inclut ou pas la dimension protection et/ou la dimension valorisation dans le champ de la politique et de l'évaluation de celle-ci. L'objet est alors davantage de mesurer une évolution, d'effectuer des bilans selon les critères des méthodes d'évaluation (impact et efficacité, efficience, etc.)

Au final, il est difficile de distinguer des connaissances qui ne seraient orientées que vers une seule des familles d'objectifs. Mais c'est dans le niveau de détail, de précisions à atteindre que la distinction peut se faire.

2. Panorama d'études et de méthodes d'étude de la fréquentation

De très nombreuses références traitent d'études de fréquentation. Elles relèvent d'une diversité parmi les objectifs traités (cf ci-dessus) et pour ce qui est de la littérature scientifique, d'une diversité d'objets de recherche. Elles sont assorties le plus souvent d'études de cas.

Les sections qui suivent (infra) mobiliseront de façon détaillée les références scientifiques pertinentes pour l'objectif premier de la tâche T1-B qui est d'estimer les effets économiques de la fréquentation du site (CPFL) par la méthode des flux monétaires issus des dépenses des visiteurs.

Concernant les références *synthétiques* orientées spécifiquement vers l'évaluation économique, des ouvrages ont déjà été cités supra (CREDOC 2008, Ministère_environnement 2017) auxquels on doit ajouter la synthèse des travaux effectués pour le compte des Grands sites de France (GSF 2008) ou des références centrées de surcroît sur certaines fréquentations, telle que celle d'événements touristiques (Min_économie and EMC 2012)

Ici il s'agit d'évoquer les références synthétiques dont le champ dépasse la stricte évaluation économique de la fréquentation, ce qui intéresse l'objectif second de Valo-sites qui est de décrire la fréquentation du site.

Pour en donner un panorama global on peut citer (EUROSTAT 2000, CRT_RA 2016).

De même une publication de l'UNESCO bien qu'orientée « évaluation économique » et « comptage de la fréquentation » dresse la liste des composantes habituelles des études de fréquentation (UNESCO 2021) :

| | | | | |
|---|--|------------------------|---------------------------------------|---|
| Trip planning | Paths / trails used | Visitor satisfaction | Local area spending | Visitor origin and travel distance |
| Importance of the protected area in choosing destination | Information / services used | Accessibility | Consumer surplus / willingness to pay | Local visitor or non-local visitor |
| Trip characteristics – party size, type of accommodations | Activities | Crowding | Health benefits | Substitute destinations and previous visits |
| Length of time spent in the local area surrounding the protected area | Amount of time spent within the protected area | Mode of Transportation | Other management-relevant questions | Demographics |

« An overview of the types of questions that are often asked in visitor surveys »

Selon cette source, les informations surlignées en vert sont celles qui sont nécessaires à l'évaluation économique (on verra que la connaissance des activités pratiquées, ou des modes de transport peuvent être également nécessaires).

Selon le guide d'observation locale du tourisme produit par le CRT RA (2016), les objectifs d'une observation des clientèles sont de connaître : la fréquentation - les caractéristiques et les comportements des clientèles - les retombées économiques. Les différents moyens de pratiquer cette observation sont parcourus.

Concernant les enquêtes « les thématiques suivantes les composent régulièrement :

- Connaître le profil des clients : âge, CSP, revenus, pays ou département d'origine
- Comprendre leurs comportements et leurs attentes : processus de choix de la destination, de réservations des différentes prestations, de comportement sur le lieu de séjour (transports, hébergements, activités pratiquées, utilisation du numérique...).
- Mesurer leur satisfaction : critères de satisfaction et d'insatisfaction, attentes... »

3. Définitions et notions-clés

a) Définitions

Un touriste est un voyageur (étranger au territoire considéré) qui passe au moins une nuit en dehors de son domicile habituel (OMT Organisation mondiale du tourisme).

Un excursionniste vient visiter le territoire considéré et retourne à son domicile dans la même journée.

Un séjour est le temps passé par le touriste dans sa destination, associé à toutes les activités pratiquées pendant ce temps.

Une excursion est le temps passé par un excursionniste dans une journée avec les activités associées. Par extension on peut parler d'une excursion que pratique un touriste partant de son lieu de séjour vers un site de visite avec retour à son hébergement temporaire dans la journée, mais il sera utile d'employer alors un terme qui garde la distinction entre l'excursion d'un touriste et l'excursion d'un excursionniste.

Une visite sera définie dans Valo-sites comme la partie d'excursion qui se déroule dans le site considéré (ici : la CPFL) : une « visite » peut comporter ainsi tous types d'activités précises (visite concrète d'un site au sein du site CPFL, mais aussi promenade, pratique d'un sport, etc.).

On parlera de « visiteurs » de la CPFL qu'ils soient touristes ou excursionnistes et quelles que soient les activités pratiquées au sein de la CPFL.

b) Géographie

Par rapport aux trois aires présentées ci-dessus, on notera que :

- Un touriste peut avoir choisi son hébergement (temporaire) au sein de la CPFL, ou au sein du territoire d'impact à étudier, ou en dehors de ce territoire (mais, par définition, à l'intérieur de l'enveloppe de toutes les destinations individuelles)
- Un excursionniste, de même, peut résider au sein de la CPFL, ou du territoire d'impact ou en dehors de ce territoire (on pourrait définir comme pour les touristes une enveloppe de tous les territoires émetteurs d'excursionnistes)

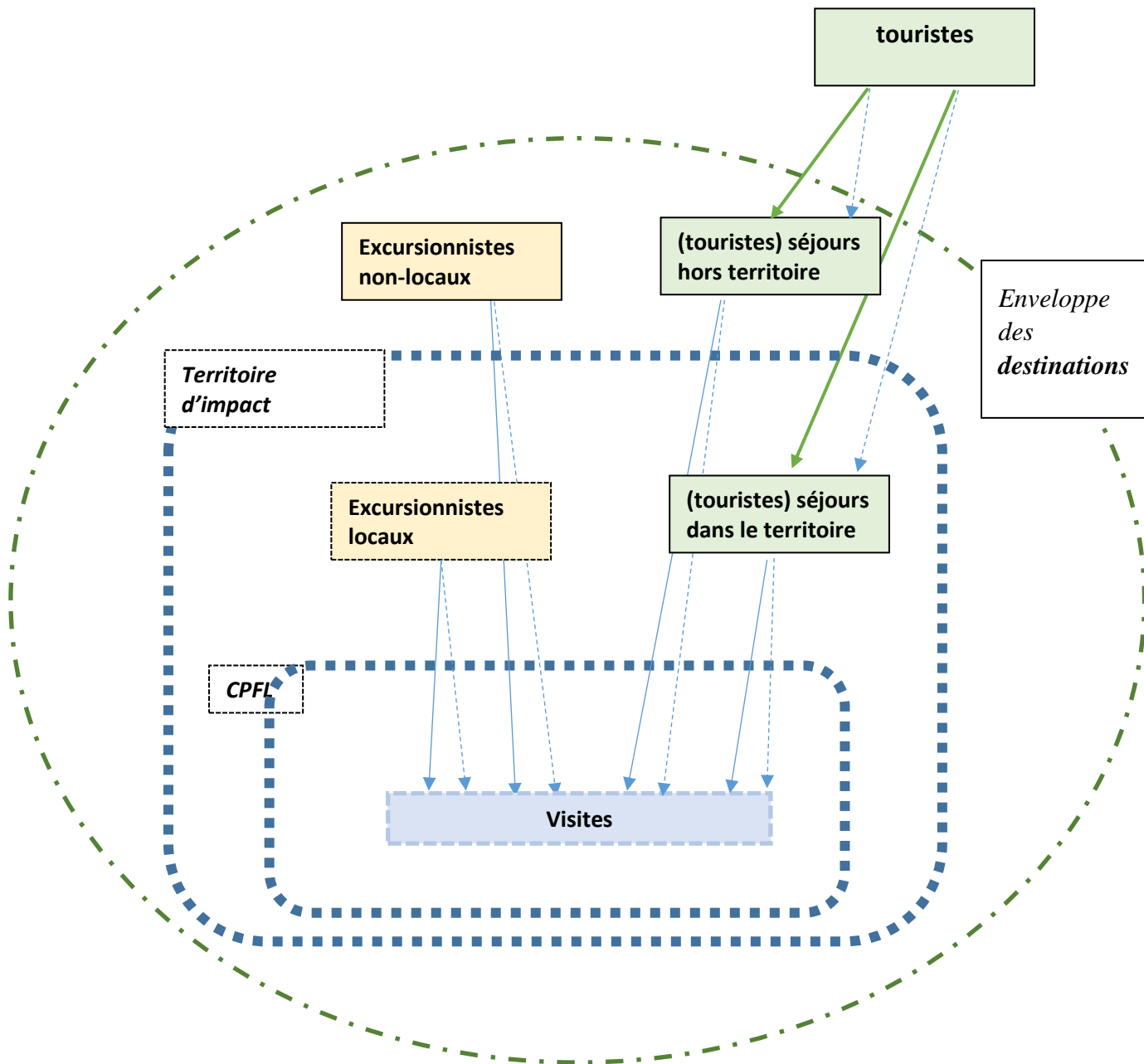
c) Occurrences

Un excursionniste, surtout s'il est proche du site, peut pratiquer plusieurs, voire de nombreuses, excursions au cours de l'année.

Un touriste peut venir plusieurs fois dans la destination au cours de l'année.

Surtout un touriste peut venir plusieurs fois sur le site au cours d'un même séjour.

Le schéma suivant résume tous ces points, les flèches pointillées rappellent les occurrences répétées possibles :



F. Evaluation des retombées économiques

En résumé de ce qui précède, la mise en place de politiques d'aménagement, de gestion et de valorisation d'un site (naturel ou non) a des objectifs multiples. Elle permet, entre autres, de préserver le site afin de le transmettre aux générations futures, de le faire connaître en le mettant en valeur afin de favoriser l'attractivité du territoire sur lequel il se situe. Le site attire un large public du fait de ses attributs. Ces politiques s'accompagnent d'investissements matériels et immatériels et sont de potentielles sources de conflit (opposition aux restrictions d'usage de l'espace par exemple). Dès lors, des informations sur les impacts économiques de ces politiques sont nécessaires afin de justifier le financement public. Aussi, la compréhension des impacts économiques des dépenses des visiteurs est utile au niveau local car les informations sur l'impact peuvent être utilisées dans le marketing, la création de nouvelles entreprises et même l'augmentation de l'acceptation générale des mesures de protection, telles que l'instauration de parcs nationaux, parmi les parties prenantes locales (Huhtala, Kajala et Vatanen, 2010).

L'augmentation des flux touristiques et résidentiels grâce à l'attractivité du territoire constitue un outil de développement local¹ puisqu'elle génère des emplois² et de revenus déterminants pour la cohésion territoriale. En effet, selon la théorie de la base économique, le développement des territoires dépend des flux monétaires externes qui viennent les irriguer (Davezies, 2000).

Ces flux proviennent du secteur basique (par opposition au secteur non basique ou domestique³) que Davezies (2004, 2009, 2010) décompose en ces quatre bases suivantes : *la base productive privée ou exportatrice* dont revenus sont issus de la production et de la vente de biens et de services en dehors du territoire considéré ; *la base résidentielle* qui comprend l'ensemble des revenus captés par les territoires grâce à leurs atouts résidentiels : retraites, salaires des migrants-alternants, dépenses des touristes ; *la base publique* composée des traitements des fonctionnaires versés par un agent extérieur : l'État) et *la base socio-médicale* constituée de revenus de transferts et remboursements de soin de santé. Cependant, l'auteur souligne que *la base résidentielle ou présenteielle* constitue avec *la base productive*⁴ les moteurs principaux des économies locales puisque la progression des bases socio-médicale et publique n'est pas une option bénéfique pour le développement des territoires même si elles restent déterminantes pour l'équilibre et la cohésion territoriale. « Le développement devient alors un enjeu de production mais aussi un enjeu de population qui capte ces revenus extérieurs » (Talandier, 2008a). Pour un territoire, note (Talandier, 2008a, 2008c, 2012), la création de richesses (économie productive), la capacité à capter les revenus de l'extérieur (économie résidentielle) ainsi

¹ Bien que l'augmentation des flux touristiques et résidentiels puisse avoir des conséquences négatives sur la société, la culture ou l'environnement (nuisance, dégradations environnementales, congestion des sites et du trafic, hausse des prix du foncier et de l'immobilier, inflation des prix locaux du logement et des prix de détail...),

² Même si l'industrie du tourisme et des loisirs est critiquée pour ces emplois saisonniers, non qualifiés et à bas salaires.

³ Le secteur basique capte les revenus à l'extérieur de la zone considérée tandis que le secteur le secteur domestique produit les biens et services demandés localement.

⁴ La base productive crée des richesses qui circulent, grâce à la redistribution vers la base résidentielle.

⁵ L'attractivité résidentielle et présenteielle étant le premier moteur du développement local des zones d'emploi françaises, au sens de la théorie de la base (Davezies, 2004).

que la maximisation de la propension locale à consommer (économie présentielle¹) constituent l'enjeu de son développement économique.

La fréquentation touristique, partie intégrante de l'économie résidentielle ou présentielle, contribue donc au développement territorial à travers l'injection de nouveaux revenus dans l'économie locale grâce aux dépenses touristiques. En tant qu'un résident temporaire, le touriste et l'excursionniste font rentrer des revenus sur un territoire et génèrent de la dépense (Pecqueur et Talandier, 2011). Ces revenus vont donc engendrer, par la suite, des effets d'entraînement sur les autres secteurs économiques locaux en termes d'emploi et de revenu. Le tourisme peut donc fournir un nouveau départ pour certains espaces ruraux économiquement dépressive (Huhtala, 2007 ; Talandier, 2008b). Les nouveaux résidents pourraient, eux aussi, stimuler une augmentation de la demande de biens et services, résultant en une économie plus diversifiée avec des emplois mieux rémunérés. En résumé, comme l'ont noté Carlino et Mills (1987), vu qu'il y a interdépendance entre la croissance de la population et de l'emploi, une prescription politique pour les responsables du développement économique local est de formuler des stratégies pour retenir ou attirer la population et l'emploi suivra.

Dans le cadre d'une analyse coûts-bénéfices², les avantages ou bénéfices procurés par la visite du site peuvent être d'ordre non marchand³, avec des effets tels que l'amélioration de la qualité du site et de l'expérience des visiteurs (bénéfices procurés aux visiteurs eux-mêmes), ou marchand avec des effets liés à l'évolution de la dépense des visiteurs et de l'offre touristique. Les auteurs d'études quantitatives sur les avantages découlant des aires de loisirs de plein air se sont concentrés sur les deux principaux bénéficiaires du tourisme et des loisirs. Ce sont, premièrement, l'utilisateur et, deuxièmement, les entreprises de l'offre et des services (McDonald et Wilks, 1986).

Il existe donc deux approches alternatives pour mesurer la valeur du tourisme et des loisirs dans les zones naturelles permettant d'interpréter son importance économique (Driml, 2010). La première approche, surplus du consommateur, est l'avantage obtenu par le consommateur lorsque ce qu'il est prêt à payer pour un bien ou un service est supérieur à ce qu'il doit payer. Il est considéré comme la mesure pertinente du bien-être économique et est fondé sur la théorie microéconomique. La seconde approche décrit les valeurs des dépenses directes associées au tourisme et aux loisirs, générant de l'activité et de l'emploi dans l'économie. Elle tient compte des dépenses des visiteurs pour les articles associés à leur voyage dans une zone, y compris, mais sans s'y limiter, le transport, les équipements de loisirs, l'hébergement et les visites commerciales. Les valeurs directes des dépenses des visiteurs entraînent également des impacts indirects résultant des achats effectués dans d'autres secteurs et des impacts induits lorsque les travailleurs dépensent leur revenu en biens et services. L'estimation de

¹ Ensemble des activités tournées vers la demande locale.

² L'analyse des coûts-avantages est également conventionnellement liée à la prise de décision associée aux propositions de projets spécifiques et est réalisée ex ante, tandis que l'analyse de l'impact économique est souvent une évaluation ex post des dépenses associées soit à un projet spécifique (achevé) soit à une activité touristique plus générale (Tyrrell et Johnston, 2016)

³ Pour estimer la valeur économique des biens et services non marchands, la théorie économique a développé plusieurs méthodes d'évaluation. Ces méthodes peuvent être classées en deux grandes catégories: les méthodes d'évaluation indirectes, dites des préférences révélées et les méthodes d'évaluation directes dites des préférences déclarées (OCDE, 2002 ; Pearce, Atkinson et Mourato, 2007 ; Brahic et Terreaux, 2010)

l'impact économique, à l'inverse de certaines méthodes d'évaluation non marchande, ne prend pas en compte les valeurs de non-usage¹. Les informations sur les impacts économiques des dépenses des visiteurs sont utiles au niveau local pour justifier le financement public et en particulier en améliorant l'acceptation générale de la politique parmi les parties prenantes locales, pour le marketing, et pour stimuler la création de nouvelles entreprises (Huhtala, Kajala et Vatanen, 2010).

1. Estimation des dépenses de la fréquentation d'un site : quels types et parts des dépenses à prendre en compte ?

Afin de connaître au mieux les retombées économiques de la fréquentation d'un site sur un territoire donné, il est primordial d'identifier les types et la part des dépenses à prendre en compte.

a) Types de dépenses à prendre à compte

Dans leur études sur la valeur économique du tourisme dans les alpes Australiennes, Mules et al. (2005) n'ont pris en compte que les dépenses des non-résidents (visiteurs à la journée et visiteurs passant au moins une nuit dans l'Etat et hors de son domicile principal) de chaque Etat dans lequel se trouvent les alpes. Les dépenses des résidents (locaux) de ces Etat ne sont pas pris compte dans l'étude d'impact économique. Cette façon de faire est d'ailleurs préconisée par bon nombre d'études d'impact économique de la fréquentation d'une attraction touristique (Bergstrom, Cordell, Watson, et al., 1990 ; Mules, 2005 ; Crompton, 2006 ; Bowker, Bergstrom et Gill, 2007 ; Huhtala, 2007 ; Crompton, 2010 ; Harnik et Crompton, 2014 ; Jeong et Crompton, 2015) ou d'un événement (McHone et Rungeling, 2000 ; Crompton, Lee et Shuster, 2001 ; Nicolas, 2007 ; Maurence, 2012 ; Warnick, Bojanic et Xu, 2012 ; Warnick, Bojanic et Cartier, 2017 ; Beckman et Traynor, 2018). L'hypothèse sous-jacente est que les dépenses des visiteurs locaux ou des résidents de la zone de l'étude d'impact représentent une remise en circulation de sommes qui existaient déjà dans la zone. En effet, les dépenses des visiteurs locaux de la zone sont perçues comme un déplacement des dépenses dans la zone d'étude plutôt que comme une demande supplémentaire pour la production de biens et de services dans la zone. Ainsi, ces dépenses liées à la visite du site ou à la participation à un événement dans la zone de l'étude d'impact se substituent en grande majorité aux dépenses en biens et services locaux qui auraient eu lieu de toute façon dans la zone.

Pour certains auteurs (Oosterhaven et van der Knijff, 1987 ; Johnson et Moore, 1993 ; Mules et Faulkner, 1996 ; Tyrrell et Johnston, 2001 ; Lindberg et Denstadli, 2004), ignorer toutes les dépenses des utilisateurs ou visiteurs locaux car représentant une recirculation d'argent déjà dans la région entraînera une sous-estimation des impacts totaux. En effet, ces auteurs pensent que de nombreux utilisateurs, en l'absence d'une attraction touristique dans leur région, iraient se récréer à l'extérieur de leur zone ou région. Leurs dépenses seraient donc perdues pour la région si les visiteurs locaux s'engageaient dans une autre activité à l'extérieur de celle-ci. Ils plaident en faveur de l'inclusion des visiteurs locaux si ces derniers auraient voyagé hors de la zone d'impact si l'attraction n'existait pas. Pour ce faire, ils préconisent la collecte d'informations supplémentaires sur leur comportement de substitution afin d'ajuster les dépenses pour refléter uniquement celles associées à la ressource

¹ Il faut noter que la valeur économique totale d'un site ou d'une ressource est constituée les valeurs d'usage et les valeurs de non-usage.

touristique d'intérêt dans la région d'impact¹. Huhtala, Kajala et Vatanen (2010) dans leur étude sur les impacts économiques locaux des dépenses des visiteurs des parcs nationaux en Finlande ont décidé d'inclure les dépenses des visiteurs locaux dans l'analyse, mais de déclarer leur contribution séparément. Snowball (2008) cité par Warnick et al. (2015) signale que le fait de ne pas définir la zone d'étude et les visiteurs locaux peut donner des résultats très différents.

Ainsi une *analyse dite d'importance*² prend en compte les dépenses de tous les visiteurs, à la fois les locaux et les non-locaux (Stynes, 1997a ; Watson et al., 2007) à la différence de l'*analyse d'impact économique qui ne prend pas en compte les dépenses des locaux*. Une analyse de l'importance est une mesure de l'importance du site (plutôt que des impacts) dans l'économie locale car elle montre la taille et la nature de l'activité économique associée aux activités récréatives / touristiques dans la région. Cependant, elle n'offre aucune information utile qui pourrait éclairer les compromis impliqués dans les décisions concernant la meilleure façon d'investir les fonds publics, contrairement à une étude d'impact (Crompton, 2006).

Comment définir la zone d'impact économique ?

La définition de la zone pour laquelle les impacts sont souhaités est l'un des éléments importants de l'étude d'impact économique. Elle permet de définir la part des dépenses des visiteurs qui sont pertinentes. Distinguer les dépenses effectuées dans la zone d'étude de celles des autres zones est particulièrement important si les visiteurs effectuent de longs trajets en s'arrêtant dans de nombreux endroits. En effet, l'enquête devrait mesurer les dépenses qui ont lieu dans la zone d'étude. Seule la partie des dépenses de voyage qui se produit dans la zone d'étude génère des impacts économiques locaux. Ces dépenses doivent être mesurées à partir du moment où le visiteur entre dans la zone d'étude jusqu'à son départ, en prenant soin d'inclure également toutes les dépenses prépayées qui reviennent aux entreprises de la zone d'étude (Stynes, 1999). Aussi, la définition de la zone d'étude segmente les visiteurs en locaux et non-locaux et affecte les multiplicateurs car plus la zone est petite, plus les multiplicateurs sont petits (Stynes, 1997b ; Huhtala, Kajala et Vatanen, 2010). La zone d'étude doit être suffisamment grande pour constituer une région économique viable. Elle devrait comprendre des installations et des sites qui pourraient être des attractions touristiques. Elle devrait également être une zone économique fonctionnelle avec une variété suffisante d'entreprises pour être raisonnablement autonome. La zone devrait également inclure les principaux marchés du travail et les zones commerçantes à une distance de trajet raisonnable (Stynes, 1997b). L'impact régional des dépenses des visiteurs dépendra de facteurs tels que les liens interentreprises au sein de l'économie régionale, la politique fiscale et la proportion du revenu local normalement dépensée dans la région. Ces paramètres dépendront eux-mêmes de la taille de la région : plus la zone est petite, moins les entreprises et les détaillants locaux sont susceptibles d'acheter des fournitures produites localement (effets indirects faibles), et moins les ménages locaux sont susceptibles d'acheter des biens produits localement (effets

¹ Il convient de noter que les visiteurs locaux pourraient probablement dépenser plus pour une activité de loisirs non locale qu'ils ne le font localement.

²Pour une présentation pédagogique de cette différenciation voir : <https://msu.edu/course/prr/840/econimpact/concepts.htm>

induits faibles)(Agnarsson, Radford et Riddington, 2008). Différentes zones d'impact sont retenues selon les études.

En France, Boncoeur et al. (2013) dans l'estimation des effets locaux de la fréquentation des sites (des îles Chausey et du Mont Saint-Michel) ont utilisé, pour des raisons conceptuelles et de disponibilité de données statistiques, les zones d'emploi (ZE) adjacentes à ces sites. Le Réseau des Grands Sites de France (2008), dans son étude sur l'évaluation des retombées économiques des grands sites de France, a choisi le niveau régional pour le Pont du Gard, niveau départemental pour le Puy de Dôme et intercommunal pour le Mont-Saint-Michel.

Aux États-Unis, les études d'impact sont réalisées au niveau comté qui constitue le plus petit niveau auquel les données économiques sont disponibles (Stynes, 1997b). Bergstrom, Cordell, Watson, et al. (1990) et Bergstrom, Cordell, Ashley, et al. (1990) dans leur étude d'impact économique des parcs ont utilisé les comtés adjacents pour former les régions d'impact local. En Finlande, les études d'impact économique de la fréquentation des parcs nationaux sont effectuées au niveau des municipalités de localisation du site et de façon discrétionnaire d'autres municipalités proches (Huhtala, 2007 ; Huhtala, Kajala et Vatanen, 2010).

En Australie, les estimations d'impact économique de la fréquentation des parcs nationaux, des parcs marins et des forêts sont produites au niveau des régions, des États ou des Territoires conformément à l'objectif de fournir des informations aux décideurs des États et des Territoires qui sont responsables du financement et de la gestion des parcs nationaux. Les études concernent alors soit un seul parc national ou un groupe de parcs. Ainsi, les impacts économiques sont estimés pour : les sous-régions Southern Forest Region et Gascoyne Coast Region de l'État d'Australie-Occidentale (Carlsen et Wood, 2004) ; les économies du Territoire de la capitale australienne, de la Nouvelle-Galles du Sud et de Victoria (Mules et al. , 2005) ; l'État de Queensland (Lindberg et Denstadli, 2004 ; Ballantyne et al., 2008) ; la région du centre de l'Australie du Territoire du Nord (Tremblay et Carson, 2007) ; la région Top End et du Territoire du Nord (Tremblay, 2007) ; l'État de la Nouvelle-Galles du Sud et les villes environnantes du parc (Mules, 2005) ; la région de Grampians de l'État de Victoria (Hughes et al., 2009).

Comment distinguer les locaux des non-locaux ?

Dans les études des retombées économiques locales de la fréquentation d'un site, il existe différentes méthodes pour distinguer les locaux des non-locaux. La méthode géographique est principalement utilisée pour les distinguer. Cette méthode géographique consiste à choisir la zone d'impact et à identifier à partir des codes postaux toute personne résidant dans cette zone d'impact comme « locale » (Lindberg et Denstadli, 2004 ; Loomis et Caughlan, 2006 ; Stynes et White, 2006 ; Bowker, Bergstrom et Gill, 2007 ; Croes et Severt, 2007 ; Loomis, 2007 ; Réseau des Grands Sites de France, 2008 ; Kemper, Popp et Miller, 2008 ; Orens et Seidl, 2009 ; Boncoeur et al., 2013 ; Beckman, Kumar et Kim, 2013 ; Kirillova, Bordelon et Pearlman, 2014 ; McGrath, Primm et Lafe, 2017). Ceux qui résident dans des zones dont le code postal est hors de la zone d'impact sont classés comme des non-locaux. Une deuxième méthode se base sur la distance parcourue de la résidence principale au site étudié. En effet, certaines études d'impact économique ont identifié les « locaux » comme les visiteurs voyageant sur une certaine distance de leur résidence principale au site étudié en aller simple (Bonn et Harrington, 2008 ; Beckman, Kumar et Kim, 2013 ; Crompton, Lee et Shuster, 2001).

Cependant, ces deux méthodes ne sont pas pleinement satisfaisantes pour différencier les locaux des non-locaux. En effet, les codes postaux sont utilisés pour simplifier et accélérer l'acheminement du courrier mais ne définissent pas comment les gens achètent réellement. Les locaux font des achats réguliers de produits et de services dans des zones de commerce ou dans des entreprises situées à proximité de leur domicile ou de leur lieu de travail (Warnick et al., 2015). D'autre part les distances parcourues offrent des moyens de distinguer différents types de visiteurs mais elles ne correspondent pas parfaitement aux définitions de « local » et « non-local ». Une nouvelle approche le *Trade Market Analysis* (TMA)¹ préconisée par l'Université du Wisconsin est appliquée dans certaines études (Warnick et al., 2011; Warnick, Bojanic et Xu, 2012, 2015; Warnick, Bojanic et Cartier, 2016, 2017, 2017; Beckman et Traynor, 2018). Cette approche utilise une combinaison des deux premières approches en plus du temps de parcourt, les zones commerciales et les attractions de la destination pour différencier les locaux des non-locaux.

Il faut noter que les non-locaux pourraient être segmentés en excursionnistes et en touristes. Contrairement aux excursionnistes qui effectuent une visite sur une seule journée mais passent la nuit dans leur résidence principale, les touristes passent au moins une nuit hors de leur résidence principale.

b) Part des dépenses à prendre à compte

Afin d'estimer avec précision les impacts économiques d'une ressource récréative spécifique, lors de l'analyse des retombées économiques de la fréquentation d'un site donné, il est nécessaire de connaître les dépenses finales qui sont associées uniquement à cette ressource ou attraction (Carlsen, 1997 ; Crompton, 2006 ; Johnson et Moore, 1993). En effet, inclure toutes les dépenses des visiteurs d'un site dans l'étude d'impact économique d'un site peut sembler problématique. Si le site n'est qu'une des nombreuses destinations ou l'une des nombreuses raisons de visiter la région, attribuer toutes les dépenses liées au voyage au site pourrait surestimer ses impacts. Loomis et Caughlan (2006) ont noté que la prise en compte de l'objectif du voyage est très importante lorsqu'il y a deux ou plusieurs attractions majeures à proximité dans la même zone géographique. Lue, Crompton et Fesenmaier (1993) ont énuméré quelques raisons qui peuvent expliquer et qui permettent de mieux comprendre les voyages à buts ou à destinations multiples. Des chercheurs ont développé des approches méthodologiques qui permettraient d'inclure les dépenses touristiques associées uniquement à la ressource étudiée. Deux approches se dégagent dans la littérature : l'approche dite d'« attribution » et celle dite de « substitution ».

Estimation générale des dépenses : approche d'attribution

Les dépenses « d'attribution » sont considérées comme la proportion de toutes les dépenses directes des visiteurs *dans la zone d'impact* qui peut être attribuée à une attraction ou à un site. Carlsen et Wood (2004)² ont utilisé un facteur d'attribution dans leur étude en se basant sur des mesures des motivations et des activités recueillies pour un échantillon de tous les visiteurs d'une région d'étude. Pour certaines études, les dépenses d'attribution comprenaient toutes les dépenses, pour tous les biens et services dans la région cible, par les touristes qui ont visité un parc national dans cette région (Lindberg et Denstadli,

¹<https://fyi.extension.wisc.edu/downtown-market-analysis/understanding-the-market/trade-area-analysis/>

² Voir aussi (Tremblay, 2007 ; Tremblay et Carson, 2007 ; Hughes et al., 2009)

2004 ; Huhtala, 2007 ; Ballantyne et al., 2008 ; Huhtala, Kajala et Vatanen, 2010). Bergstrom, Cordell, Watson, et al. (1990) et Cordell, Bergstrom et Watson (1992) ont affecté toutes les dépenses engagées par les non-résidents sur le site ou à proximité immédiate du site à la zone d'impact. Aussi, une partie des dépenses annuelles hors domicile et hors la zone d'impact en équipement durable ou autres a été allouée à la zone d'impact local.

Buultjens et Luckie (2004) ont multiplié les dépenses quotidiennes moyennes totales par personne par : un facteur proportionnel de 1 si la visite du parc était le seul objectif du voyage et par le rapport entre le nombre de jours passés dans le parc divisé par le nombre de jours de vacances ou par 0,5 si la visite fait partie d'un voyage multi objectif. Le Réseau des Grands Sites de France (2008) dans ses études sur le Pont du Gard et le Puy de Dôme a défini des taux d'affectation, à partir de certaines questions sur la motivation de présence, allant de 0% à 100% applicables aux dépenses. Les dépenses journalières des excursionnistes sont prises en compte à 100%. De leur côté, Boncoeur et al. (2013) ont appliqué pour les visiteurs dont le déplacement hors domicile principal est spécialement motivé par la visite du site toutes les dépenses dans la zone d'impact considérée et seulement les dépenses sur le site des visiteurs dont le déplacement hors domicile principal n'est pas spécialement motivé par la visite du site. Pour saisir l'influence relative d'une attraction dans le contexte plus large des attractions cumulatives¹ dans une région donnée, Jeong, Crompton et Dudensing (2016) préconise d'incorporer dans le questionnaire une question du type : « Sur une échelle de 0 à 10 indiquez le nombre qui représente le mieux la mesure dans laquelle la visite du parc était le principal objectif de votre voyage dans cette région² » au lieu d'utiliser des questions dichotomiques. Ainsi, les dépenses de chaque répondant seront multipliées par leur mesure de proportionnalité.

Estimation prudente des dépenses : approche de substitution

Les dépenses de « substitution » sont considérées comme le montant des dépenses qui auraient été perdues s'il n'y avait pas le site (ressource ou attraction) ou si ce dernier n'était pas disponible pour les visiteurs. Ainsi, seules les dépenses des non-locaux qui n'auraient pas visité la région (ou des locaux qui auraient voyagé en dehors de la région) s'il n'y avait pas le site sont incluses dans l'estimation de l'impact (Johnson et Moore, 1993 ; Carlsen et Wood, 2004 ; Lindberg et Denstadli, 2004 ; Tremblay, 2007 ; Tremblay et Carson, 2007 ; Nicolas, 2007 ; Ballantyne et al., 2008 ; Driml et McLennan, 2010). Il s'agit en fait d'une estimation du montant des dépenses touristiques supplémentaires qui peuvent être associées à l'existence du site. En d'autres termes, il s'agit d'une estimation de la réduction des dépenses touristiques dans le contexte hypothétique et contrefactuel où les visiteurs n'auraient **pas** accès au site. Cette approche consiste à fournir une estimation prudente (conservatrice) des dépenses des visiteurs. Cela donne une indication des revenus réels attirés dans une région par le site étudié. Ainsi, pour Tyrrell et Johnston (2016), l'approche fondamentale consiste à mesurer la différence en termes économiques entre deux situations: une situation réelle (ou le statu quo) et une situation hypothétique

¹ Ce terme indique qu'un nombre donné d'attractions dont le principal marché cible est les touristes accroîtront leurs activités si elles sont situées sur un même itinéraire, à proximité ou dans un ordre logique les unes par rapport aux autres, que si elles sont largement dispersées.

² Voir (Tyrrell et Johnston, 2001 ; Martínez-Espiñeira et Amoako-Tuffour, 2009) pour d'autres applications de cette méthode.

(les scénarios avec et sans). Pour les impacts économiques associés à une attraction touristique, la comparaison pourrait être les impacts avec (la situation réelle) et sans (la situation hypothétique) la présence de l'attraction. Dans le cas d'une ressource naturelle, la substitution mesure la volonté des touristes de substituer un voyage dans la zone d'impact à un voyage en dehors de la zone d'impact, si la ressource naturelle n'existe pas (Jones et Wood, 2008).

Pour identifier le facteur de substitution, Carlsen et Wood, 2004, Tremblay, 2007 et Tremblay et Carson, 2007 ont conçu une question de scénario examinant le choix probable de destination de vacances des répondants si le parc n'existait pas. C'est la proportion de tous les visiteurs de la région qui ne viendraient pas qui est appliquée au montant des dépenses d'attribution. Contrairement à ces auteurs, c'est pour les visiteurs qui ont effectivement visité le parc, que Ballantyne et al. (2008), ont estimé les dépenses de substitution, qu'ils appellent les dépenses « générées par les parcs nationaux » en appliquant au montant dépensé dans la région par ces visiteurs, la proportion de ceux qui ne viendraient dans la région si le parc n'existait pas.

Dans leur étude sur les retombées économiques liées au tourisme dans des parcs nationaux, Job (2008) et Mayer et al. (2010) ont étudié l'affinité des visiteurs envers le parc qu'ils visitent, à travers des questions successives (connaissance du statut de protection du site et visite du site était la principale raison de venir dans la région). Ils ont ensuite ajusté les valeurs calculées avec la part ou pourcentage des visiteurs ayant une affinité élevée pour les parcs nationaux. Huhtala, Kajala et Vatanen (2010) dans leur étude sur l'estimation des retombées économiques locales du tourisme dans les parcs nationaux, ont inséré dans leur questionnaire une question concernant l'importance du parc national dans le choix de la destination. Seules les dépenses de ceux qui avaient répondu que le parc national était leur seule ou la plus importante destination du voyage sont retenues pour le calcul de l'impact proprement dit.

Lindberg et Denstadli (2004) ont défini leur mesure prudente en ne tenant compte que des dépenses qui ont été faites dans la région par les visiteurs mais qui n'auraient pas été faites si le parc n'existait pas. Ce montant comprend les dépenses des touristes qui ne viendraient pas du tout en l'absence du parc. Tisdell et Wilson (2002) et Stoekli, Smith et Newsome (2005) utilisent une question hypothétique sur ce que les visiteurs auraient fait si les attractions étudiées n'existaient pas. Seules les dépenses des visiteurs qui n'auraient pas du tout voyagé dans la région si l'attraction n'existaient pas et une partie des dépenses de ceux qui auraient passé moins de temps / moins de jours sont prises en compte.

2. Mesure de l'activité économique

Les responsables du financement, de la planification et de la gestion des parcs et d'autres grands espaces ouverts s'intéressent de plus en plus aux informations sur l'activité économique du tourisme et des loisirs dans la zone ou espaces environnants où se situent ces parcs ou espaces ouverts. Dans le contexte d'une analyse économique régionale, l'activité économique mesure les dépenses réelles et leur cycle dans l'économie de la région. On peut distinguer entre autres *l'importance ou contribution économique et l'impact économique* (Watson et al., 2007). *Une contribution économique* est définie comme les changements bruts dans l'économie existante d'une région qui peuvent être attribués à une industrie, un événement ou une politique donnée. Elle mesure tous les revenus associés à une industrie ou à une activité, même si toute l'activité économique ne représente pas une nouvelle richesse pour une économie. *Cette mesure est utile en particulier pour tous les acteurs (au sens large, incluant les*

politiques publiques) de cette activité. Ainsi, l'effet direct dans une analyse de contribution inclut les achats par les consommateurs locaux et non locaux.

En revanche, *une analyse d'impact économique* est généralement basée sur l'activité économique créée uniquement à partir d'une nouvelle richesse (Bangsund et Leistritz, 2003). *Les impacts économiques* sont définis comme les changements nets dans une économie à la suite d'un événement spécial, d'un festival ou d'une politique. L'impact économique peut se produire soit en raison de l'attraction de dépenses des non-résidents dans l'économie, soit en réduisant les fuites d'argent des résidents de l'économie (Long et Perdue, 1990). Les effets directs dans une analyse d'impact économique consiste en la partie de la demande finale des revenus générés par une industrie à partir des ventes hors de la région (exportations), la portion marginale de la demande finale des visiteurs qui viennent dans la région (dépenses touristiques) et les dépenses marginales de la demande finale des locaux qui auraient dépensé leur argent pour des biens en dehors de la région si l'industrie n'avait pas été présente dans la région (substitution des importations) (Watson et al., 2007).

Les études d'impact économique régional sont précieuses pour quantifier les contributions tangibles des loisirs et du tourisme à l'activité économique locale. Elles prévoient des changements dans la production totale des secteurs, le revenu des ménages et l'emploi découlant des dépenses des visiteurs. Ces trois éléments constituent généralement les indicateurs de mesure de l'impact économique (McDonald et Wilks, 1986). L'analyse d'impact économique de la fréquentation d'un site touristique et de loisirs suit les flux d'activité économique associés au site afin d'identifier les changements dans les ventes, les recettes fiscales, les revenus et les emplois dans la région dus aux opérations de gestion du site ou aux dépenses des visiteurs (c'est-à-dire, comment l'argent dépensé par les visiteurs se manifeste dans l'économie locale).

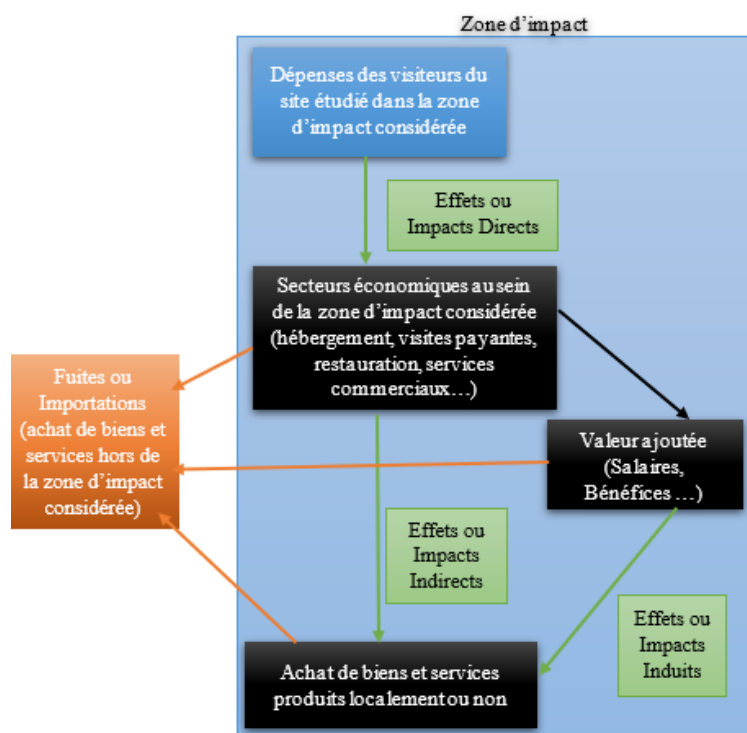


Figure 1 : Détermination des effets directs, indirects et induits

Ce schéma montre l'analyse de l'impact économique standard. Il retrace le chemin emprunté par l'argent une fois qu'il quitte la poche du visiteur.

Les flèches vertes représentent les effets directs, indirects et induits.

- ✚ **Les Effets directs** constituent le premier flux d'argent provenant des dépenses des visiteurs vers les entreprises auxquelles ils paient directement (e.g. commerces de détail, hôtellerie-restauration).
- ✚ **Les Effets Indirects** se manifestent lorsque des entreprises qui servent principalement les visiteurs achètent des produits et des services à d'autres entreprises employant des populations locales.
- ✚ **Les Effets Induits** se produisent lorsque les ménages (main-d'œuvre), qui ont perçu le salaire supplémentaire, en retour, achètent des biens de consommation dans l'économie locale (consommation finale).

Les flèches orange représentent **les fuites**. Une fuite se produit lorsque l'argent s'échappe de l'économie d'une région parce qu'un consommateur local (ménage ou entreprise) a acheté un produit auprès d'un fournisseur extérieur.

Distinction entre effets :

Les évolutions de fréquentation peuvent créer des *modifications de l'offre*, de façon plus complexe (effets de seuil, etc.)

Les effets *indirects* de la fréquentation (voir Tâches 2 et 4.2) *créent eux-mêmes des effets induits*,

On peut ainsi **envisager d'autres effets** (de plus long terme) :

- Directs, par modification de l'offre
- Induits, supplémentaires, par le jeu des effets indirects

Dans cette étude seront estimés **les seuls effets induits qui découlent des effets directs immédiats**.

Distinction entre territoires d'impact emboîtés

Par ailleurs, il est indispensable, dans le cadre de l'estimation de l'impact économique, d'éviter l'agrégation inappropriée des effets. En effet, lorsque l'unité géographique d'analyse est modifiée, cela change la définition des visiteurs. Par exemple, prenons quatre départements A ; B ; C et D formant la région R. L'impact économique de la fréquentation d'un site peut être estimé distinctement pour les quatre départements ou pour la région.

Cependant, cet impact économique sur R n'est pas égal à la somme des impacts économiques sur A ; B ; C et D. Puisqu'en utilisant non pas les départements mais la région comme unité géographique d'analyse, tous les visiteurs qui résident les quatre départements deviennent automatiquement des visiteurs locaux du point de vu de la région. Ainsi donc les dépenses de ces derniers ne sont pas prises en compte dans le cadre de l'impact économique sur la région. Les seules dépenses qui peuvent être incluses dans cette analyse est celle qui sont engagées par les visiteurs qui vivent hors de la région R. En revanche, cette agrégation peut être appropriée si toutes les dépenses, quel que soit le type de visiteurs, sont prises en compte. C'est le cas d'une analyse de la contribution économique.

a) Estimation des effets directs ou primaires

Les effets directs constituent le premier flux d'argent provenant des dépenses touristiques vers les entreprises et les agences gouvernementales auxquelles les touristes paient directement (Stynes, 1997a). Ces dépenses touristiques, correspondant à la consommation touristique totale, sont estimées généralement à partir d'enquêtes socioéconomiques en face à face et/ou en auto-administration auprès des visiteurs du site étudié. Ces enquêtes permettent aussi de collecter des informations sur le comportement et le profil des visiteurs (durée de séjour, lieu de résidence, activité pratiquée, motivations, etc...). L'hébergement, la restauration, les transports, les divertissements et le commerce de détail sont les secteurs touristiques primaires directement touchés par ces changements. Le montant des dépenses totales des visiteurs sont obtenues en multipliant la dépense moyenne¹ par visiteur par jour par la durée moyenne du séjour et le nombre total de visiteurs (Stynes, 1997b ; Carlsen et Wood, 2004 ; Mules et al., 2005 ; Wood et al., 2006 ; Tremblay, 2007 ; Prideaux et Falco-Mammone, 2007 ; Tremblay et Carson, 2007 ; Ballantyne et al., 2008 ; Hughes et al., 2009).

Pour améliorer la précision des résultats, il est recommandé de diviser les visiteurs en groupes en fonction de leur type d'hébergement ou d'autres facteurs affectant leurs dépenses (Stynes, 1997b, 2002 ; Stynes et Sun, 2003 ; Stynes, 2007 ; Driml et McLennan, 2010). Des segments devraient être formés pour regrouper les visiteurs ayant des habitudes de dépenses similaires, réduisant ainsi la variance des moyennes de dépenses et la taille d'échantillon nécessaire pour les segments individuels (Stynes, 1999). Il est aussi recommandé de calculer les dépenses moyennes par segment touristique sur la base d'échantillons d'au moins 50 à 100 visiteurs par segment touristique puisque les dépenses peuvent varier considérablement selon le type de touriste. Les effets directs sont donc calculés pour chaque catégorie de dépenses en multipliant les dépenses moyennes par visiteur dans chaque groupe de visiteurs par le nombre de visiteurs dans ce groupe. Les estimations de l'impact économique reposent sur de bonnes estimations du nombre et des types de visiteurs (Stynes, 1997b). Pour réduire les variances des moyennes de dépenses, Jeong, Crompton et Dudensing (2016) préconise de calculer séparément les dépenses des visiteurs de jour et de nuit car ils ont des habitudes de dépenses différentes.

¹ Gazel et Schwer (1997) et Buultjens et Luckie (2004) ont utilisé la valeur médiane comme la meilleure estimation des dépenses quotidiennes moyennes dans leur étude car ils pensent que les répartitions des dépenses ont tendance à être biaisées positivement (la moyenne est affectée de manière significative par les observations individuelles de dépenses élevées). Cependant, la dépense médiane dans les enquêtes auprès des visiteurs est généralement inférieure à la moitié de la dépense moyenne (Stynes et White, 2006).

En résumé, le calcul des effets directs passe par les étapes suivantes : calcul des dépenses moyennes par visiteur dans la zone d'impact considérée ; calcul des dépenses totales des visiteurs dans la zone d'impact considérée et la conversion ou non des dépenses directes en valeur ajoutée, emplois et de revenus.

(1) Calcul des dépenses moyennes par visiteur dans la zone d'impact considérée

Deux moyennes d'échantillons peuvent être calculées : les dépenses moyennes par personne ou les dépenses moyennes par personne par jour. Les dépenses moyennes par personne (c'est-à-dire pendant tout leur séjour dans la zone d'impact considérée peuvent ensuite être multipliées par les données démographiques sur le nombre total de visiteurs dans la région) pour obtenir les dépenses totales. Les dépenses moyennes par personne et par jour sont utilisées si les données sont en jours-visiteurs (Driml et McLennan, 2010). Les dépenses moyennes par personne par jour peuvent être calculées en utilisant une pondération individuelle ou une pondération de groupe (Jeong, Crompton et Dudensing, 2016) :

La pondération individuelle divise le total des dépenses totales des groupes par le nombre total de jours-visiteurs :

$$\frac{\sum_{i=1}^n E_i}{\sum_{i=1}^n (T_i D_i)} = \bar{e}$$

Où E_i représente les dépenses totales de chaque groupe i , T_i est la taille du groupe pour chaque groupe (nombre de personnes), D_i est la durée de séjour de chaque groupe (nombre de jours), n est le nombre total de groupes ou de réponses dans l'échantillon, et \bar{e} est la moyenne pondérée des dépenses par personne pour l'échantillon.

Si la pondération de groupe est utilisée, les dépenses moyennes par personne et par jour de chaque groupe sont d'abord calculées :

$$\frac{E_i}{T_i D_i} = e_i$$

Où E_i , T_i et D_i sont définis comme ci-dessus et e_i représente les dépenses moyennes par personne et par jour pour le groupe i . Les dépenses quotidiennes moyennes par personne pour chaque groupe sont ensuite moyennées à nouveau dans l'échantillon, créant une moyenne d'une moyenne :

$$\frac{\sum_{i=1}^n e_i}{n} = \acute{e}$$

Où \acute{e} représente les dépenses moyennes par personne pondérées par groupe pour l'échantillon.

Les auteurs ont constaté en analysant leurs données que l'utilisation de pondérations de groupe, plutôt que de pondérations individuelles, est susceptible de produire des dépenses moyennes considérablement exagérées¹.

(2) Calcul des dépenses totales des visiteurs dans la zone d'impact considérée

Pour estimer la dépense totale qui être appliquée à un modèle d'impact économique ou à un ensemble de multiplicateurs pour estimer l'impact économique direct, l'approche habituelle de la grande majorité des études d'impact sur les loisirs et le tourisme consiste à multiplier le nombre de visiteurs soit par la dépense moyenne, soit par la dépense moyenne par jour multipliée par la durée moyenne du séjour.

Dépense totale = dépense moyenne par visiteur par jour x durée moyenne de séjour² x nombre total de visiteurs.

Impact économique

Dépenses totales = **dépense moyenne d'attribution** (des non-locaux) **ou de substitution** (des non-locaux et des locaux) par visiteur par jour x durée moyenne du séjour x nombre total de visiteurs.

Importance économique

Dépenses totales = **dépense moyenne d'attribution** de tous les visiteurs (locaux et non-locaux) x durée moyenne du séjour x nombre total de visiteurs

b) Estimation des effets secondaires

Les effets secondaires (indirects et induits) peuvent être estimés en termes de production brute, de ventes, de revenus, d'emploi ou de valeur ajoutée (Stynes, 1997b). Les effets indirects se manifestent lorsque des entreprises qui servent principalement des touristes achètent des produits et des services à d'autres entreprises employant des populations locales. Les impacts induits se produisent lorsque les ménages (main-d'œuvre), qui ont perçu le salaire supplémentaire, eux aussi en retour, achètent des biens de consommation dans l'économie locale (Johnson et Moore, 1993 ; Frechtling, 1994 ; Stynes,

¹ Sun et Stynes (2006) préconise aussi la pondération individuelle qu'ils appellent « estimateur de ratio » pour estimer les dépenses moyennes.

² Calcul de la durée moyenne du séjour : Moyenne pondérée ?

$$\bar{D} = \frac{\sum_{i=1}^n T_i D_i}{\sum_{i=1}^n T_i}$$

1997b ; Henry et Deane, 1997 ; Stynes, 2005 ; Bowker, Bergstrom et Gill, 2007 ; Kemper, Popp et Miller, 2008 ; Miller et Blair, 2009 ; Huhtala, Kajala et Vatanen, 2010). Lorsque des produits ou services intermédiaires sont achetés à l'extérieur de la région, une partie des revenus liés au tourisme fuit à l'extérieur de la région. Ces fuites doivent être soustraites de l'augmentation du chiffre d'affaires pour calculer les effets indirects (Huhtala, 2007). L'ampleur de ces effets secondaires est donc directement liée à la propension des entreprises et des ménages locaux à acheter à des fournisseurs locaux (Stynes, 1997b).

Approche prédominante adoptée pour mesurer les effets secondaires de la fréquentation sur des territoires (Frechtling et Horváth, 1999 ; Dwyer, Forsyth et Spurr, 2004 ; Mayer et al., 2010), les modèles d'entrées-sorties, mis en évidence par Léontief en 1951, sont actuellement la méthode la plus courante pour les études d'impact économique du tourisme. Ils intègrent les interdépendances sectorielles dans une économie et permettent de désagréger l'impact économique des dépenses touristiques entre les secteurs (Fletcher, 1989 ; McDonald et Wilks, 1986). Les tableaux d'entrées-sorties (TES), appelés également tables ou analyses Input-Output, sont utilisés dans un grand nombre d'études pour mesurer les effets secondaires de la fréquentation sur des territoires variés (e.g. Archer & Fletcher 1996 ; Zhou et al. 1997 ; Stynes 1999 ; Dwyer et al. 2004a ; Lindberg & Denstadli 2004 ; Powell & Chalmers 2006 ; Zhang et al. 2007 ; Huhtala et al. 2010).

Aux Etats-Unis, l'estimation des effets secondaires à l'échelle locale se fait à partir d'algorithmes et de bases de données au niveau comté grâce au logiciel IMPLAN (Bergstrom, Cordell, Ashley, et al., 1990 ; Cordell, Bergstrom et Watson, 1992 ; Johnson et Moore, 1993 ; English et al., 1996 ; Hjerpe et Kim, 2007). Contrairement aux Etats-Unis, en France et dans de nombreux pays européens, les tableaux entrées-sorties ne sont pas disponibles au niveau locale (Huhtala, 2007 ; Mayer et al., 2010 ; Boncoeur et al., 2013). Aussi, les modèles d'entrées-sorties ont plusieurs limites. Ils supposent que les ressources, telles que la main-d'œuvre, les terres et les capitaux, circulent librement vers le tourisme et les industries connexes. Ces ressources sont effectivement supposées ne pas être utilisées ailleurs ; ils ne proviennent pas d'autres industries et n'entraînent pas de réduction de la production ailleurs. De cette façon, les techniques d'entrées-sorties et de multiplicateurs comptent les influences positives sur l'activité économique, mais ignorent les influences négatives (Dwyer, Forsyth et Spurr, 2004 ; Kumar et Hussain, 2014).

Pour améliorer les modèles d'entrées-sorties, d'autres modèles sont proposés : les matrices de comptabilité sociale (MCS) et les modèles d'équilibre général calculables (EGC) (Dwyer, Forsyth et Spurr, 2004 ; Loveridge, 2004). Les MCS fonctionnent avec le même ensemble de base d'hypothèses et de méthode de résolution que les modèles entrées-sorties, mais accordent plus d'attention aux aspects distributionnels des chocs modélisés, de sorte qu'ils sont utilisés lorsqu'une attention particulière est accordée au développement économique, par opposition à une simple croissance économique. Les modèles CGE sont susceptibles d'être plus petits car l'ajout de contraintes d'offre et de demande réduit l'impact des nouvelles dépenses des visiteurs, contrairement aux hypothèses d'offre parfaitement élastiques dans les modèles d'entrées-sorties (Loveridge, 2004).

G. Etude de la fréquentation d'un site : aspects méthodologiques détaillés

1. Méthode d'échantillonnage

Pour recueillir des données sur les caractéristiques des visiteurs et des voyages, telles que les dépenses effectuées, le but de la visite, les lieux visités et les activités de loisirs entreprises dans un site et d'autres données démographiques, la collecte de données primaires est requise s'il n'y a pas de données secondaires fiables disponibles. En effet, il est important de vérifier l'existence de données secondaires avant d'entreprendre toute collecte de données. Les données secondaires ont l'avantage d'être immédiatement disponibles et généralement à un coût minimal. Cependant, ces données ont plusieurs inconvénients, par exemple, elles fournissent généralement des informations limitées et spécifiques qui ne répondent pas aux besoins de la recherche car collectées à d'autres fins (Eurostat, 2000).

Pour recueillir des informations sur les caractéristiques d'une population donnée, l'idéal serait de collecter des données sur toutes les composantes de cette population. Dans le cadre d'une étude visant une grande population, une telle collecte de données peut coûter en ressources disponibles (en termes d'argent, de personnel et de temps). La solution pour réduire ces coûts consiste à tirer une conclusion sur les caractéristiques de cette population en analysant une partie (échantillon) réduite et représentative de cette population. La théorie des sondages est un ensemble d'outils statistiques permettant l'étude d'une population au moyen de l'examen d'une partie de celle-ci (Tillé, 1999). Elle se consacre essentiellement au problème de la sélection de l'échantillon et à la recherche d'estimateurs.

Un sondage désigne donc une méthode statistique d'analyse d'une population-mère à partir d'un échantillon d'individus prélevés au sein de cette population-mère. Il s'oppose donc au recensement qui est l'étude exhaustive de la population-mère. Il existe deux grandes catégories d'échantillonnage : l'échantillonnage probabiliste (ou aléatoire)¹ et l'échantillonnage non probabiliste (ou empirique ou à choix raisonné)²³. Dans le cas du premier, chaque unité a une chance connue (non nulle) d'être sélectionnée. Grâce à la randomisation (la sélection au hasard ou aléatoire) qui est une caractéristique du processus de sélection, la variabilité de l'échantillonnage est estimée et le biais possible identifié.

Dans le second cas, chaque unité n'a pas une probabilité de sélection connue (la probabilité peut même être nulle). La distribution des caractéristiques à l'intérieur de la population étant supposée égale. En effet, les échantillons visent à être représentatifs par construction, à partir de critères sur lesquels on peut viser à ce qu'ils aient les mêmes caractéristiques que la population mère. Il peut donc exister un biais dit de sélection et les marges d'erreur peuvent être difficile à calculer. Pour pouvoir procéder à l'échantillonnage probabiliste, il faut généralement disposer d'une base de sondage (nombre et caractéristiques des unités composant la population-mère). Mais il peut arriver qu'il n'existe pas de

¹ Échantillonnage aléatoire simple ; échantillonnage systématique ; échantillonnage avec probabilité proportionnelle à la taille ; échantillonnage stratifié ; échantillonnage en grappes ; échantillonnage à plusieurs degrés ; échantillonnage à plusieurs phases.

² Échantillonnage de commodité ou à l'aveuglette ; échantillonnage volontaire ; échantillonnage au jugé ; échantillonnage par quotas ; échantillonnage sur place

³ <https://www150.statcan.gc.ca/n1/edu/power-pouvoir/ch13/nonprob/5214898-fra.htm>

base de sondage ou qu'il soit difficile ou coûteux d'en disposer. Ceci conduit souvent à l'utilisation de l'échantillonnage non probabiliste.

a) La population cible

Elle est définie au regard des objectifs de l'étude. Dans le cadre d'une étude économique de la fréquentation d'un site naturel, il est important d'enquêter toute personne visitant le site. Ainsi, la population de l'étude correspond donc aux individus visitant le site et exerçant une ou des activités payantes et/ou gratuites (sentiers de randonnées, châteaux, lacs aménagés, sites de loisirs...). Cette population est composée de non-locaux (touristes¹ et visiteurs à la journée) et des locaux (résidents). Par locaux ou résidents, on entend les visiteurs du site étudié habitant la zone d'impact considérée. Les autres visiteurs étant des non-locaux. La distinction entre ces deux types de population est importante car les non-locaux et les locaux ont généralement un impact différent, à la fois en termes économiques, environnementaux et sociaux, sur l'économie locale. Le suivi de ces deux types de visiteurs influence le choix de l'enquête à réaliser et son organisation.

b) Choix du lieu d'enquête

Dans le cadre de la collecte de données primaires, une fois la population cible identifiée, il est important de choisir le(s) lieu(x) où les enquêtes vont être réalisées. Ce choix se fait en fonction de la zone d'étude (fermée ou ouverte²) et des informations à collecter. Selon Eurostat (2000) quatre lieux d'enquête alternatifs et / ou complémentaires peuvent être identifiés : les points d'entrée / sortie de la zone d'étude, à bord des moyens de transport³, les établissements d'hébergement des visiteurs et les sites touristiques visités ou leurs points d'accès (parkings, dont certains sont communs à plusieurs sites).

Tableau 1 : Le lieu d'étude suggéré en fonction du type de zone

| Lieu de l'enquête | Zone fermée | | | Zone ouverte | | |
|------------------------------|-------------|---------------|------------|------------------|-----------------|------------|
| | Grand pays | Petite Île | Attraction | Grande région | Petite Ville | Attraction |
| Aux points d'entrée / sortie | Oui | Oui | Non | Non | Non | Non |

¹ Visiteurs qui passent au moins une nuit dans la zone considérée en dehors de leur domicile principal

² Zone dans laquelle il n'existe aucun mécanisme de contrôle de l'entrée des visiteurs.

³ Il est nécessaire d'obtenir la coopération des membres d'équipage pour la distribution de questionnaires aux passagers.

| | | | | | |
|---------------------------------------|-----|-----|-----|-----|-----|
| A bord des moyens de transport | Oui | Oui | Non | Non | Non |
| Dans les établissements d'hébergement | Non | Oui | Non | Oui | Non |
| Sites touristiques visités | Non | Oui | Oui | Oui | Oui |

Tableau 2 : Le modèle d'enquête proposé en fonction du type d'informations à collecter

| Information | Zone fermée | | | Zone ouverte | | |
|--------------------------------------|-------------|-------------------------|------------|----------------------|----------------------|------------|
| | Grand pays | Petite Île | Attraction | Grande région | Petite Ville | Attraction |
| Flux de visiteurs | Points E/S | Points E/S ou Transport | Sites | Sites ou hébergement | Sites ou hébergement | Sites |
| Caractéristiques des visiteurs | Points E/S | Points E/S ou Transport | Sites | Sites ou hébergement | Sites ou hébergement | Sites |
| Qualité de l'expérience de la visite | Points E/S | Points E/S ou Transport | Sites | Sites ou hébergement | Sites ou hébergement | Sites |
| Comportement des dépenses | Points E/S | Points E/S ou Transport | Sites | Sites ou hébergement | Sites ou hébergement | Sites |

Point E/S = Points d'entrée/sortie ; Transport = à bord des moyens de transport ; Sites = Sites touristiques visités ; hébergement = dans les établissements d'hébergement

Source : (Eurostat, 2000)

c) Difficulté ou impossibilité pratique de création d'une base de sondage

Bien souvent lors de la préparation d'une enquête, on ne dispose pas de base de sondage permettant de sélectionner les individus dans une liste donnant un accès direct aux unités de la population. C'est le cas des populations rares¹ ou mobiles² (Kalton, 1991, 2001 ; Marpsat et Razafindratsima, 2010) telles que la population touristique. En effet, la population touristique est dispersée et a une taille qui n'est pas connue ou qui varie dans le temps.

Pour enquêter la population touristique, Eurostat (2000) préconise l'échantillonnage aléatoire à plusieurs degrés comme méthode la plus efficace. Il consiste, après avoir à identifier deux ou plusieurs populations mutuellement exclusives respectant une sorte d'ordre hiérarchique, à échantillonner d'abord une première unité de sondage, puis une deuxième unité à l'intérieur de la première et à poursuivre le processus jusqu'à l'obtention d'un échantillon final. L'échantillonnage aléatoire à plusieurs degrés est généralement stratifié au premier stade, de manière aléatoire ou systématique (Cochran, 1977) si l'objectif est d'obtenir des informations détaillées sur les caractéristiques des visiteurs, du voyage (durée du séjour, mode de transport utilisé, etc.) et sur leur comportement de consommation. On parlera donc de l'échantillonnage aléatoire stratifié³ à plusieurs degrés.

Manente et Minghetti, (1998) ont appliqué la technique d'échantillonnage aléatoire à deux degrés lors d'une enquête sur les visiteurs de la zone archéologique de Pompéi et d'Herculanum. Dans cette étude, la période de sept jours, correspondant à une semaine d'enquête a été tirée au hasard de chaque mois (de mai à novembre 1997) en tenant compte des périodes de pointe et des périodes de soudure, pour un total de 49 jours d'échantillonnage. Ce qui constitue les unités de la première étape. Les personnes visitant chaque site archéologique ou l'église étaient les unités de la deuxième étape. Les visiteurs ont été enquêtés en mettant en œuvre un tirage systématique.

Deville et Maumy (2005, 2006) ont utilisé cette approche lors d'une enquête portant sur la fréquentation touristique d'origine intra ou extrarégionale en Bretagne. Leur base de sondage est constituée de trois strates : les achats en boulangerie, les visites d'un ensemble de sites emblématiques de la Bretagne, le passage au péage autoroutier de « La Gravelle ». Dans chaque strate un échantillonnage à plusieurs degrés est effectué. Ainsi, dans la première strate « les achats en boulangerie » un échantillon à 3 degrés est effectué (boulangeries, de jours d'enquête et de clients dans la boulangerie à un jour donné). Dans la deuxième strate « les visites d'un ensemble de sites emblématiques de la Bretagne » un échantillon à 2 degrés est effectué (de jours d'enquête, de personnes qui passent sur un des 16 sites référés à un jour donné). Enfin dans la troisième strate un

¹ Personnes ayant connu le placement dans leur enfance, personnes à très hauts revenus ou patrimoine, sans-domicile, usagers de drogues injectables, personne socialement stigmatisée (prostitution), etc.

² Nomades, sans-domicile, Visiteurs de bibliothèques et des parcs, électeurs dans les bureaux de scrutin, personnes qui font leurs courses, les malades en consultation externe, les voyageurs et les passagers des automobiles, touristes, etc.

³ Le principe de la stratification consiste à diviser la population étudiée en groupes homogènes (strates) selon un critère lié à la ou les variables d'intérêt et à réaliser un sondage dans chacune d'elles.

échantillon à 2 degrés est effectué (jours d'enquête et de personnes qui passent au péage autoroutier de La Gravelle à un jour donné).

Cette méthodologie est utilisée par l'INED et INSEE pour réaliser des enquêtes auprès des sans-domicile (Ardilly et Le Blanc, 2001 ; Brousse, Guiot de la Rochère et Massé, 2006 ; Marpsat, 2008).

De façon générale, cette méthode consiste notamment à sélectionner les lieux produisant des services à destination des touristes après en avoir fait une liste exhaustive, puis tirer un échantillon des touristes qui les fréquentent en tenant compte des moments où ces derniers fréquentent ces lieux. Ainsi, l'échantillonnage lieux-moments (Kalton, 1991 ; Marpsat et Razafindratsima, 2010) consiste, dans un premier temps, à établir une liste complète des sites fréquentés et des moments (ou jours) où ils sont fréquentés ; ensuite, à tirer les couples (site-jour d'enquête) et à enquêter des visiteurs sur le site sélectionné au jour sélectionné. Ces couples constituent les unités primaires d'échantillonnage (UPE) de la base de sondage. On peut constituer des strates à partir de la liste des UPE en tenant compte des caractéristiques des sites ou de la saisonnalité de leur fréquentation. Cette méthode est préconisée par English et al. (2002) pour estimer les visites dans toutes les forêts nationales et utilisée par Bowker, Bergstrom et Gill (2007) dans leur étude sur l'estimation de la valeur économique et des impacts des sentiers récréatifs (Virginia Creeper Rail Trail). Dès lors que les UPE sont échantillonnées, un échantillon de visiteur est tiré habituellement par échantillonnage aléatoire ou systématique dans ces UPE. Pour réaliser cet échantillonnage systématique, il faut cibler les endroits où les visiteurs peuvent entrer et sortir du site visité. Si l'échantillonnage systématique proprement dit ne peut être réalisé pour une raison ou une autre (augmentation du nombre de visiteurs le jour de l'enquête par exemple), une méthode proposée par consiste à prendre un échantillon systématique d'intervalle de temps et de choisir le prochain visiteur qui entre (ou qui sort) après chaque d'intervalle de temps échantillonnée. Une autre méthode consiste à choisir la $n^{\text{ième}}$ personne à entrer (ou à sortir) après la fin de la dernière interview, n peut avoir comme valeur 2, 3, 4, ou 5. Il faut bien sûr tenir compte de la nature du flux (caractéristique clé de la population mobile). Si le flux entre les UPE est uniforme, les UPE peuvent être échantillonnées avec une probabilité égale de visites. Dans le cas contraire, une probabilité proportionnelle à la taille (estimée si la taille n'est pas connue) est requise. Toutefois, il est à noter que ces méthodes produisent un échantillon non probabiliste, avec le risque de biais que cela entraîne. Pour la majorité des enquêtes sur les flux de populations humaines, la visite, plutôt que le visiteur, est l'unité d'analyse appropriée (Kalton, 1991). Un des inconvénients de cette méthode est qu'on peut se confronter à un biais de couverture². Ce biais est dû par exemple, à la non fréquentation des sites d'enquête par une forte proportion des unités faisant partie de la population étudiée (Marpsat et Razafindratsima, 2010).

Ces approches peuvent poser un autre problème lors de l'estimation des caractéristiques de la population étudiée si les liens entre les unités des deux populations ne sont pas bijectifs. En effet, un visiteur peut fréquenter plusieurs sites touristiques de la zone étude lors de son séjour (ou dans la période de l'étude). Ce dernier a donc une probabilité élevée d'être sélectionné. Il est donc primordial

¹ Parfois appelé échantillonnage par intervalles, l'échantillonnage systématique consiste à choisir les éléments en respectant un intervalle fixe entre les tirages, le point de départ ayant été d'abord déterminé au hasard.

² L'erreur de couverture apparaît quand la population dans laquelle on prend l'échantillon ne "couvre" pas la totalité de la population que l'on veut étudier.

de tenir compte de ce fait dans les estimations. Pour qu'une variable d'intérêt donnée soit estimée sans biais à partir des valeurs issues de l'enquête des visiteurs, il suffirait de pondérer chaque valeur par le poids d'estimation (inverse de la probabilité de sélection ou d'inclusion) de chaque visiteur. Or dans le cas de l'échantillonnage des visiteurs d'un espace ouvert sans point de contrôle, le calcul de ces probabilités est généralement difficile, voire impossible. Afin de résoudre ce problème, Eurostat (2000) juge nécessaire d'estimer la probabilité que le visiteur soit interrogé au moyen d'une enquête complémentaire ou une enquête préliminaire. Certains auteurs préconisent dans ce cas utilisation de la méthode généralisée du partage des poids (MGPP)¹ (Lavallée, 1995 ; Ardilly et Le Blanc, 2001 ; Leroy, 2018). La MGPP² permet d'obtenir un poids d'estimation pour chaque visiteur enquêté. Elle permet une pondération correcte mais pas forcément de façon optimale. Ce poids est calculé par exemple à partir du poids de sondage des couples (site-jour d'enquête) sélectionnés et du nombre de sites visités par chaque visiteur enquêté (liens). Etant donné qu'un visiteur a une probabilité élevée d'être enquêté si le nombre de sites qu'il a visités est élevé, il est par conséquent logique que son poids issu de la méthode de partage des poids diminue. Comme le souligne Loonis (2009), ce qui importe dans la méthode du partage des poids, c'est le nombre de fois où un individu aurait pu être sélectionné, ce qui est très différent du nombre de fois où il a effectivement été sélectionné.

d) Détermination de la taille de l'échantillon

La détermination de la taille optimale de l'échantillon dépend de la méthode de sélection de l'échantillon choisie. Le choix de la taille dépend avant tout du compromis entre la précision et des ressources disponibles³. Ces deux éléments jouent un rôle fondamental. On cherchera à minimiser les coûts une fois la précision de l'estimation fixée. Dans tous les cas, on doit s'assurer du niveau maximum de précision compte tenu des ressources limitées disponibles. Plus la taille de l'échantillon est grande plus l'analyse sera précise. Mais il est à préciser que la taille de l'échantillon n'est pas en lien (ou très peu) avec la taille de la population mère. Si la taille de la population cible est connue, la formule est :

$$n = \frac{Np(1-p)}{\frac{(N-1)e^2}{t^2} + p(1-p)} = \frac{t^2 Np(1-p)}{(N-1)e^2 + t^2 p(1-p)} \quad (2)$$

Lorsque cette dernière n'est pas connue, la formule devient :

$$n = \frac{t^2 p(1-p)}{e^2} \quad (1)$$

Avec :

¹ Elle est utilisée dans 3 cas : sondage indirect, échantillon sélectionnée dans plusieurs bases de sondage non-disjointes et échantillons totalement partiellement panélisés.

² Pour plus de précision sur cette méthode, il faut consulter : (Lavallée, 1995 ; Ardilly et Le Blanc, 2001 ; Deville et Maumy, 2005 ; Deville et Lavallée, 2006 ; Deville et Maumy, 2006)

³ Argent, de personnel et de temps

N = taille de la population

n = taille de l'échantillon attendu.

p = proportion estimative de la population présentant la caractéristique étudiée dans l'étude. Lorsque cette proportion est ignorée, une pré-étude peut être réalisée ou sinon $p = 0,5$ sera retenue.
e = marge d'erreur (traditionnellement fixée à 5%)

t = niveau de confiance déduit du taux de confiance (traditionnellement 1,96 pour un taux de confiance de 95%) – loi normale centrée réduite.

2. Traitement des non-réponses et redressement d'échantillon

a) Traitement des non-réponses

La non-réponse peut se définir comme l'incapacité à obtenir des réponses utilisables, pour tout ou partie des questions de l'enquête. Elle constitue l'une des sources d'erreur¹ dans les sondages. Cette erreur est subie et non maîtrisée. En effet, ceux qui ne répondent pas peuvent différer systématiquement des répondants et l'enquête sera par conséquent biaisée. Lors des entretiens en face-à-face, les raisons les plus courantes de non-réponse sont le refus de l'enquêté échantillonné de répondre à certaines questions ou son impossibilité de le faire en raison du manque d'information (Eurostat, 2000). On distingue la non-réponse totale et la non-réponse partielle. Il y a non-réponse totale quand certains individus échantillonnés ne peuvent finalement pas être enquêtés ou si l'ensemble du questionnaire ou une trop grande partie du questionnaire est inexploitable. Ce type de non-réponse peut être corrigé par la repondération ou par imputation (Statistique Canada, 2003 ; Caron, 2005 ; Deroyon, 2018). On parle de la non-réponse partielle lorsque le répondant a omis de répondre à certaines questions ou si seules certaines questions sont inexploitables (Statistique Canada, 2003 ; Deroyon et Favre-Martinoz, 2018). Ces données manquantes peuvent, lors de la phase de traitement des données, entraîner une distorsion de l'estimation de la variable que la question manquante représente. Dès lors, une méthode efficace et peu coûteuse doit être utilisée pour combler ces lacunes. La non-réponse partielle est corrigée par imputation.

Il existe donc deux techniques de correction de la non-réponse : la repondération et l'imputation. La repondération consiste à augmenter le poids d'échantillonnage des répondants pour tenir compte des non-répondants. Il existe plusieurs méthodes de repondération² dont la méthode des groupes de réponse homogène et le calage sur marges sont les plus utilisées dans les enquêtes de la statistique publique en France (Deroyon, 2018). L'imputation vise à remplacer les réponses manquantes par des valeurs plausibles. Dans le détail, en ce qui concerne les variables qualitatives, il n'existe généralement aucune autre information susceptible de remplacer les données manquantes. Cependant, il n'y a aucune raison de penser que les caractéristiques des non-répondants soient très différentes de celles des

¹ Erreur de mesure, biais de sélection, biais de couverture de la population

² Pour plus d'approfondissement de ces méthodes, voir notamment (Statistique Canada, 2003 ; Brilhault et Caron, 2004 ; Caron, 2005)

répondants (Eurostat 2000). La méthode Hot-deck (Statistique Canada, 2003 ; Vanderschelden, 2005) peut être utilisé pour imputer les variables qualitatives. En ce qui concerne les variables quantitatives, plusieurs méthodes d'imputation des valeurs manquantes sont utilisées (Statistique Canada, 2003). On peut citer, entre autres, le colddeck, l'imputation par la moyenne, l'imputation par la médiane, l'imputation par le ratio, l'imputation par la régression, l'imputation tendance unitaire, l'imputation par le plus proche voisin (Statistique Canada, 2003 ; Brilhault et Caron, 2004 ; Caron, 2005). Ainsi, la valeur manquante peut être remplacée, soit en tirant une autre réponse, pour la même question au hasard, soit en utilisant une valeur moyenne ou médiane obtenue à partir de partir de l'échantillon total ou d'un sous-échantillon particulier qui présente les mêmes caractéristiques que celles du visiteur dont la non-réponse est examinée, soit en utilisant un modèle de régression pour prédire la variable dont les valeurs sont manquantes.

b) Le redressement des pondérations

Lors du dépouillement d'une enquête (prétraitement des informations collectées lors de l'enquête), il arrive qu'on constate que les données collectées ne sont pas représentatives de la population étudiée. En effet, selon Tillé (1999), il devrait avoir autant que possible une ressemblance entre un échantillon et la population à étudier de sorte que certaines catégories apparaissent en même proportion dans l'échantillon et dans la population. L'exactitude des résultats dépend souvent de la représentativité de l'échantillon de la population. Cette représentativité se fait sur les caractéristiques pouvant influencer la ou les variables d'intérêt (faisant l'objet de l'étude). Dans une étude sur les dépenses touristiques liées à la fréquentation d'un site par exemple, les variables auxiliaires possibles qui peuvent être corrélées à la variable d'intérêt (dépense) sont : type d'hébergement, effectif du groupe de voyage, durée du séjour, types de visiteurs (touriste ou excursionniste).

Les enquêtes statistiques analysent des échantillons d'une population afin de révéler les caractéristiques de la population générale. Dans la pratique, certaines caractéristiques ne sont pas prises en compte lors de l'échantillonnage ou même sont difficiles à utiliser sur le terrain lors de l'enquête et sont donc ignorées pour ne pas compliquer le travail de l'enquêteur. Il arrive, dès lors, que les proportions des données collectées ne correspondent pas aux proportions attendues (celles de la population-mère). Les résultats de l'enquête peuvent être améliorés en intégrant des informations auxiliaires¹ (supplémentaires) dans le calcul des estimateurs. Ces informations auxiliaires sont toute information connue sur la population soit à travers un recensement ou soit tout simplement par la base de sondage. Elles peuvent être : le total d'un caractère sur la population, des sous-totaux selon des sous-populations, des moyennes, des proportions, des variances, les valeurs d'un caractère sur toutes les unités de la base de sondage (Tillé, 1999).

Le processus d'amélioration des résultats est généralement appelé redressement de l'échantillon. Le redressement a pour but de corriger l'échantillon enquêté pour qu'il soit à l'image de la population-mère (Lemel, 1976). Le redressement consiste à affecter des poids aux réponses des enquêtés dans

¹ Il faut noter que ces informations peuvent être utilisées soit au niveau du tirage de l'échantillon, soit au niveau de l'expression de l'estimateur. Si elles sont nombreuses, une partie peut servir à améliorer le tirage, et l'autre à améliorer l'estimateur (Ardilly, 2006)

l'analyse des résultats de telle sorte que l'on retrouve alors les proportions connues sur la population. Ce sont les poids de sondage des individus de l'échantillon (et, par prolongement, les estimateurs) qui sont redressés, et non pas l'échantillon (Ardilly, 2006). Il est à noter qu'avant de passer au redressement de l'échantillon, il faut effectuer un test d'ajustement ou d'adéquation de test du Khi-2 qui compare globalement la distribution observée dans un échantillon statistique à une distribution théorique.

A part le redressement des pondérations, on a le redressement par suppression ou par extraction d'un échantillon représentatif a posteriori. Ce type de redressement se fait en supprimant ou en extrayant aléatoirement des répondants parmi les catégories surreprésentées après la collecte des données. L'objectif est de retrouver les proportions attendues (celles de la population-mère). Cette méthode a pour inconvénient de réduire la taille de l'échantillon (données collectées).

La méthode de redressement des pondérations diffère de la précédente car elle ne conduit pas à la réduction de la taille de l'échantillon. Dans un échantillon représentatif de la population, **une unité statistique représente d'un certain nombre d'unités statistiques non sondées et de profils similaires et se voit attribuer un poids de sondage.** Lors d'une étude on peut constater, une fois l'enquête terminée, que l'échantillon recueilli n'est pas représentatif de la population étudiée. Le redressement des pondérations vise donc à réajuster ces poids de sondage initiaux selon des critères définis.

Les méthodes des pondérations impliquent une pondération des unités dans un échantillon, de sorte que les valeurs de certaines variables correspondent aux valeurs externes « fiables », telles que les informations recueillies lors d'un recensement. Le principe général de la méthode est que, une fois que l'échantillon reproduit à l'identique certains traits de la population (précédemment connus), on suppose que les variables d'intérêt reproduiront également par la suite étroitement leurs valeurs « réelles » (inconnues) au sein de la population générale (Barbu, 2010). Les informations auxiliaires sont utilisées comme critères de redressement à travers des poids de sondage qui subissent des changements, qui à leur tour entraînent des changements d'estimateur. Les différentes méthodes de redressement¹ des pondérations sont (Roy et Vanheuverzwyn, 2001 ; Ardilly, 2006) :

- La post-stratification ou la stratification à postériori : lorsqu'on désire redresser l'échantillon sur une ou plusieurs variables auxiliaires qualitatives ;
- L'estimation par le quotient ou estimation par ratio : lorsqu'on désire redresser l'échantillon sur une seule variable auxiliaire quantitative ;
- L'estimation par la régression : lorsque nous avons plusieurs variables auxiliaires quantitatives ;
- Le calage sur marge, soit la généralisation des trois méthodes ci-dessus

Dans le cas de la post-stratification, si on souhaite utiliser au moins deux variables auxiliaires et si on dispose de l'effectif de chacun des croisements de variables auxiliaires dans l'échantillon et dans la population, il suffit d'utiliser la formule ci-dessus. Si en revanche, on connaît les effectifs des croisements dans l'échantillon, mais on ne dispose que sur la population, des marges ou des effectifs

¹ <https://lemakistatheux.wordpress.com/category/collecte-des-donnees-ponderation-gestion-de-la-non-reponse-et-transformation-des-donnees/le-redressement-des-ponderations-dun-echantillon/>

par catégories sur chacune des variables prise isolément, on utilisera l'algorithme « du raking ratio¹ » pour effectuer le redressement.

Il est important de vérifier la corrélation entre la ou les variables auxiliaires et les principales variables d'intérêt. En effet, les variables auxiliaires doivent être fortement corrélées au sujet de l'enquête. Toutes les informations disponibles ne peuvent donc pas être utilisées. Les informations auxiliaires peuvent provenir d'un recensement, d'une enquête précédente ou d'une enquête auprès d'un plus grand échantillon. Ces informations peuvent être intégrées dans le plan d'enquête et / ou peuvent être utilisées pour équilibrer l'échantillon une fois extrait, afin de fournir une meilleure qualité des résultats de l'enquête (Barbu, 2010).

3. Traitement des voyages et séjours à forfait

Il arrive que la personne enquêtée fasse partie d'un voyage ou séjour à forfait (voyage organisé). Ce visiteur achète donc une combinaison de plusieurs types de services pour le même voyage ou les mêmes vacances. Cette combinaison peut être faite à l'avance par un voyageur ou une agence de voyage ou peut être davantage personnalisé avant de la conclusion du contrat. Elle peut comprendre les services comme : le transport, l'hébergement, la location d'une voiture, la restauration, les excursions locales ou tout autre service touristique.

À des fins statistiques et analytiques, il arrive qu'on veuille connaître les dépenses par service consommé. Souvent les visiteurs ne savent pas la part du budget qui revient à chaque poste de consommation. Il se pose alors un problème de ventilation des dépenses. Ce problème peut être résolu de la meilleure façon en enquêtant directement les voyageurs et les agences de voyages (tant résidentes que non-résidentes opérant dans le pays). Dans le cas où cette approche n'est pas réalisable, une répartition proportionnelle du forfait entre les différents services le composant peut être fait sur la base des dépenses engagées pour ces mêmes services par d'autres visiteurs individuels n'ayant pas acheté de forfait. Le poids de chaque service est calculé sur les dépenses globales de chaque visiteur, et ces pourcentages sont appliqués au forfait. Pour augmenter la précision des résultats, les visiteurs individuels doivent faire partie des sous-groupes homogènes selon certaines caractéristiques.

¹ Voir Ardilly, 2006 pour plus de détail

III. Le cas de la CPFL : l'enquête socio-économique auprès des visiteurs

A. La Chaîne des Puys – faille de Limagne

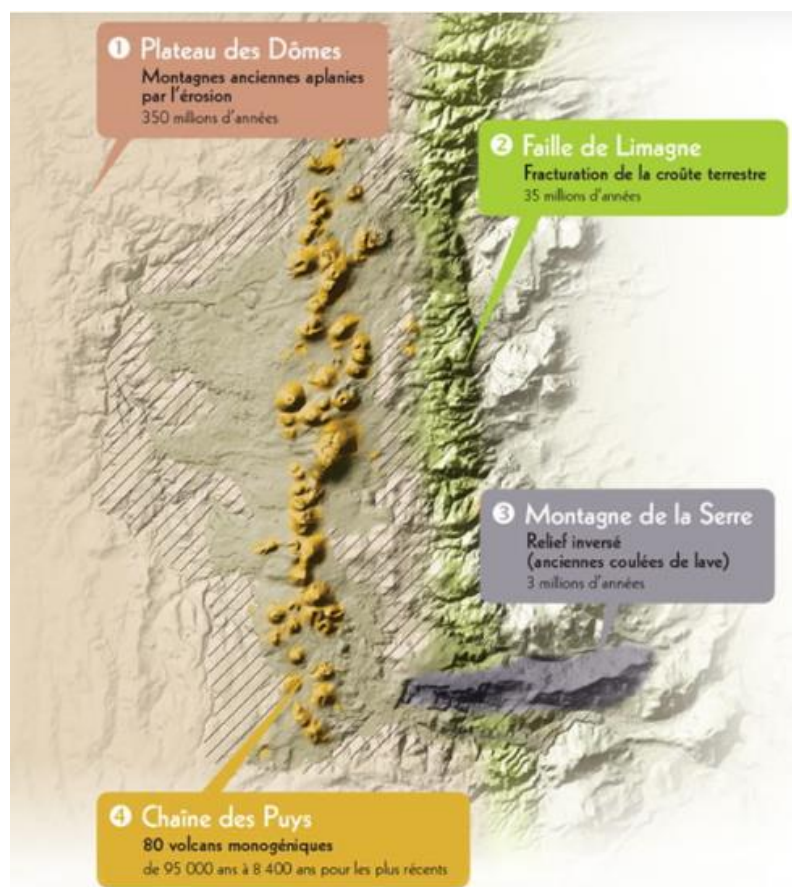


Figure 2 : Chaîne des Puys-Faille de Limagne

Source : Conseil départemental du Puy-de-Dôme

B. Conception et réalisation de l'enquête

1. Le type d'enquête

L'objectif de la première phase de cette tâche 1-B du projet est de réaliser un « état des lieux » de la fréquentation aussi complet que possible. Parmi les diverses formes d'enquête possibles (en face à face ou auto-administrée que ce soit sous forme papier ou en ligne) il apparaissait que l'enquête en face à face était la mieux à même de recueillir le plus d'informations possible.

Par ailleurs le calendrier du projet ne permettait pas de mener une enquête préalable consistante permettant de tester toutes les facettes de l'exercice, mais l'enquête en face à face permettait un retour des enquêteurs et une mise au point du protocole dans un premier temps.

Concernant les périodes à couvrir, sachant que des fréquentations diverses de la CPFL se produisent en toutes saisons, mais avec des différences entre saisons, il a été décidé d'assurer la couverture de l'année entière (ce que la crise du COVID n'a pas permis totalement).

Concernant les lieux d'enquête, la Chaîne des Puys Faille de Limagne (CPFL) étant une zone ouverte et les visiteurs enquêtés étant des touristes et des excursionnistes, la réalisation d'une enquête aux points d'entrée /sortie ou à bord des moyens de transport est écartée. Une enquête dans les établissements d'hébergement ne couvrirait que les touristes. De plus, les touristes qui visitent la CPFL ne sont hébergés sur le site que dans une proportion très minoritaire au vu des capacités d'hébergement sur le site, et on ne connaît pas, avant enquête, l'étendue de la zone concernée par cette fonction d'hébergement des touristes visitant le site. Enfin une enquête dans les établissements d'hébergement ne couvrira pas tous les types de logement (échapperont domiciles d'amis, de parents ou les résidences secondaires).

Le type d'enquête « sur les sites touristiques visités » est donc très utile dans notre cas. Il consiste à interroger les personnes sur leur lieu d'activité de loisir, c'est-à-dire l'endroit où l'on a une probabilité de retrouver la population recherchée. Elle permet d'enquêter les visiteurs étrangers et nationaux d'une part et d'obtenir des informations précises sur les touristes et les excursionnistes de l'autre. En outre, ce type d'enquête peut être utilisé pour évaluer la pression touristique qui affecte les sites et les régions environnantes.

2. Questionnement et questionnaire

L'information centrale recherchée est celle du montant de dépenses des visiteurs, assortie de : 1- la part de ces dépenses qui bénéficient au territoire choisi, donc la localisation des dépenses, 2- la distinction entre visiteurs selon leur origine, et 3- la part des dépenses qui est attribuable à l'existence et aux qualités de la CPFL (« l'effet CPFL »).

Ainsi le cœur du questionnement est de connaître, sur les visiteurs :

- Leurs Profils (touristes ou excursionnistes, nature et taille du groupe, nationalité, âge, CSP)
- Leurs Comportements (hébergement choisi, activités menées, moyens de transport) et La géographie de la fréquentation : (trajets, lieux d'hébergement, lieux d'activités)
- Leurs Motivations et part de la CPFL dans leur choix de destination (« effet CPFL »)
- Les Dépenses détaillées du jour ou du séjour.

Sur cette base un questionnaire a été construit comportant un cœur de questions couvrant les thèmes ci-dessus et des questions optionnelles, posées si les circonstances le permettaient.

Une version synthétique (thématique) du questionnaire est donnée ci-après, sachant que la version opérationnelle, implantée sur tablette, avec ses dynamiques de contrôle à la saisie, de branchements conditionnels selon les réponses, est portée en annexe.

Le questionnaire a été soumis pour avis aux partenaires du projet dont certains ont proposé l'ajout d'une question (cas du PNR), il a été testé en Mai et Juin 2019, il a évolué pendant cette période.

a) **Le questionnaire (version synthétique, par thèmes) :**

Partie A : Nature de séjour et de visite

- + Touriste ou excursionniste ?
- + Hébergement (Résidence principale, secondaire, hébergement temporaire)
- + Mode de tourisme : (fixe, itinérance)

Partie B : Visite de la CPFL

- + Précision sur le contour du site (avec carte simplifiée)
- + Activités de la journée
- + Fréquence des visites du site
- + Connaissance antérieure du site

Partie C : Patrimonialisation et UNESCO

- + Connaissance du « label » PNR, concernant le site
- + Connaissance de l'appellation UNESCO patrimoine mondial et de l'inscription CPFL
- + Influence des « labellisations » sur le choix de visite
- + Influence des « labellisations » sur choix de destination

Partie D : Choix de la destination et perceptions (Touristes)

- + Choix entre destinations possibles et motifs
- + Echelle d'influence de la Chaîne des Puy dans le choix de destination
- + Quelle décision si la CP se trouvait ailleurs ?

Partie F : Activités du séjour

- + Toutes activités pratiquées au cours du séjour (en CPFL ou ailleurs)

Partie G : Hébergement

- + type d'hébergement
- + lieu d'hébergement
- + durée du séjour

Partie H : Dépenses

- + Effectif concerné par le budget (adultes, enfants)
- + Budgets/dépenses :

Les questions sont posées une fois pour le jour de visite, et une fois pour les dépenses globales du séjour

- hébergement
- restauration (restaurants et cafés)
- alimentation (autre que restauration)
- achat de souvenirs et de produits locaux
- visites, événements, activités sportives, culturelles, et achats d'équipement
- transport en destination (modes de déplacements, et budget)
- services à la personne
- + Ou :
 - budget global, si l'enquête ne veut ou ne peut pas détailler

Partie J : Informations générales

- + *genre, âge, catégorie socioprofessionnelle*
- + *lieu de résidence (pays, département, commune (si en 63)), nationalité*
- + *configuration : seul, en couple, en famille, en groupe.*
- + *Effectifs du groupe (adultes/enfants)*
- + *qui est à l'initiative de la visite*

Partie K : Et après ?

- + *recommanderiez-vous ce site à vos proches ?*
- + *avis sur la visite de la Chaîne des Puyx Faille de Limagne*
- + *intention de revenir dans les deux prochaines années ?*
- + *avis sur votre séjour*
- + *avis sur les aménagements du site*
- + *(autres) remarques, suggestions*

Partie L : Compléments optionnels

Cette partie n'est administrée que si le temps le permet / si la personne souhaite encore répondre.

- + *définition spontanée du site*
- + *perception de la destination*
- + *précisions sur connaissance CP, Unesco, GSF*
- + *motivations pour venir sur le site*
- + *motivations pour venir dans la destination*
- + *conscience de la présence d'autres activités humaines et des risques de perturbation*

Après entretien d'enquête:

- + *Remarque sur le déroulement du questionnaire*
(c'est l'enquêteur qui indique toute réflexion utile pour la suite)

b) Précisions :

Les réponses sur les activités pratiquées (du jour et du séjour) étaient notées sur une liste de choix pré-renseignée, comportant un choix « autre », liste qui s'est allongée en cours d'enquête pour faciliter les saisies. Cette liste constituait une aide à la saisie et également à la relance des enquêtés. Chaque activité pré-renseignée était doublée d'une case de commentaires en saisie libre qui permettait de noter les détails (nom de Puy, nom de lieu, d'activité atypique, etc.)

Les réponses sur les dépenses étaient notées, pour chaque type de dépense, sous deux formes au choix : une liste de plages de valeurs (pour les enquêtés peu précis), une case de saisie en valeur numérique, assorties d'une case de commentaires.

c) Limites du questionnaire

Certaines questions n'ont pas été posées, ce qui crée des incertitudes dans les phases de comptage, d'analyse et d'estimation des effets économiques. Elles concernent :

- les activités dénombrables pratiquées par les excursionnistes répétés en dehors du jour d'enquête, (et les budgets estimés). On verra plus loin qu'une hypothèse a été faite de 50% de visites dénombrables.
- les séjours répétés des touristes concernés (leur nombre, leur taux de visite de la CPFL dans l'année, les activités dénombrables au cours de ces séjours, les budgets estimés).

Avoir simplifié la fréquentation des touristes en ignorant que certains séjours sont le fait des mêmes individus ne modifie pas nécessairement les estimations économiques mais biaise les études des profils des touristes. Le biais est limité par le fait que les touristes « répétés » seraient environ 11% du total, selon le SDT 2018 établi par Kantar (ex-TNS Sofres) pour le département du Puy de Dôme. Si la fréquentation de la CPFL n'est pas un motif habituel de séjour répété au cours d'une même année (hypothèse), alors ce taux sera inférieur à 11% pour les touristes fréquentant la CPFL.

3. Protocole d'enquête

a) La base de sondage

Dans le cadre de notre enquête, la base de sondage est constituée d'une liste d'unités primaires d'échantillonnage (UPE) fondée sur les seules variables secondaires disponibles, *faute d'autre connaissance préalable sur la population des visiteurs*, à savoir : les périodes de fréquentation et les sites touristiques visités ou leurs points d'accès (parkings, dont certains sont communs à plusieurs sites).

Le principe est donc de faire en sorte que la répartition quantitative des enquêtes suive autant que possible la répartition quantitative des fréquentations par UPE : période X lieu.

(1) L'information préalable, partielle et indirecte sur la fréquentation

La non-disposition de données secondaires sur la fréquentation totale, période par période ou non, de la CPFL, pour façonner l'échantillonnage quantitatif des enquêtes, a conduit à tester au sein d'une autre tâche du projet, l'application d'autres méthodes d'estimation de la fréquentation totale comme « *Flux-vision* » de l'opérateur Orange et le comptage routier. En attendant les conclusions de ces tests, on peut s'appuyer sur les nombres de visites dites « dénombrables » telles qu'elles sont données partiellement par les sites disposant de billetteries (les trois entreprises Panoramique, Vulcania et Lemptégy incluant la Maison de la Pierre) ou de dispositifs de comptage (Maison du PNR, pavillon des sources de la Société des eaux de Volvic) et aussi les données des « éco-compteurs ». Les éco-compteurs équipent les sentiers à la fois les plus fragiles et les plus fréquentés qui mènent aux sommets des Puy les plus connus.

Ces sources de dénombrement ne couvrent pas les fréquentations sans visite payante, et hors des sentiers qui mènent à ces sommets. C'est pourquoi on parle d'information *partielle*.

Il faut noter que ces données cumulées indiquent un nombre de visites mais pas un nombre de visiteurs puisqu'un même visiteur peut pratiquer la visite de plusieurs sites au cours de son excursion ou de son séjour. Cette question sera traitée plus loin avant de procéder aux totalisations de données telles que les dépenses des visiteurs. C'est pourquoi on parle d'information *indirecte* sur la fréquentation.

Le parti pris dans l'ensemble de la tâche présentée ici est de s'assurer que les hypothèses et les choix de valeurs (corrections, imputations), toutes choses égales par ailleurs, ne risquent pas de surestimer les résultats, tant en matière de fréquentations que de résultats économiques en aval de la fréquentation.

(a) Propositions autour de l'utilisation des éco-compteurs

Le principe de comptage des éco-compteurs est simple. A chaque fois qu'une personne passe devant le capteur de ce dernier elle est comptabilisée en fonction de son sens de passage (IN pour les personnes qui montent et OUT pour celles qui descendent).

La question qui se pose est celle de la conversion entre le nombre de passages devant tous les éco-compteurs relatifs à ce Puy et le nombre de visiteurs de ce Puy. En faisant l'hypothèse sous-jacente qu'un visiteur d'un puy fait un aller-retour devant un seul capteur, ou un seul « aller » devant deux éco-compteur du site s'ils sont plusieurs et disposés en séquence, les données mensuelles ou annuelles ont été divisées par 2 pour avoir la fréquentation de chaque puy. On rappelle que c'est une « estimation-plancher » qui est recherchée, or cette hypothèse peut surestimer la fréquentation. En effet, un visiteur pour accéder à un puy peut passer devant plusieurs éco-compteurs ou même pour une raison inconnue faire plusieurs aller-retours devant le même éco-compteur, mais ce cas est supposé très minoritaire (cas d'aller-retours répétés). De plus, à l'inverse il faut noter que l'utilisation des données issues des « éco-compteurs » sous-estime le nombre de visiteurs réels car certains sentiers n'ont pas d'éco-compteurs. Il est donc nécessaire d'affiner ces données avec la participation d'acteurs qui connaissent bien les sites.

Les cas complexes sont ceux de Puys (ou un couple de Puys comme ceux de La Vache et de Lassolas) qui disposent de deux voire trois éco-compteurs : on peut faire l'inventaire de tous les circuits possibles (certains sentiers permettent aussi de ne pas passer devant un éco-compteur). Seule une petite partie de ces circuits fait passer le visiteur devant tous les éco-compteurs du site.

Enfin certains Puys ne disposent pas d'éco-compteurs ce qui exclut certains visiteurs de la CPFL, et certains Puys avec éco-compteurs offrent des circuits qui ne passent pas du tout ou pas deux fois devant un éco-compteur. De plus il est possible que certains visiteurs (par exemple des enfants) passent à coté de l'éco-compteur, même si globalement ils suivent l'itinéraire officiel.

Exemple de cas complexe : Puys de La vache et Lassolas :
trois éco-compteurs (EC) et au moins un sentier non équipé. Exemples d'itinéraires possibles :

Itinéraire 1 : 1*EC1 puis 1*EC2 fin de boucle courte : 2 détections

Itinéraire 2 : 1*EC1 puis 1* EC3 fin de boucle longue : 2 détections

Itinéraire 3 : 1* EC1 puis sortie par sentier non équipé : 1 détection, etc...

Donc on relève au maximum 2 détection par visiteur, sauf le cas supposé rare :

Itinéraire 4 : 1* EC1 puis 1*EC2 puis remontée devant le même EC : 1*EC2 puis 1*EC3 boucle avec deux pétales : 4 détections

On voit qu'ici diviser le nombre de passages par 2 donnera un nombre de visiteurs sans doute approché (l'itinéraire 3 joue dans un sens et le 4 dans l'autre sens), ce qui est une façon de rester dans l'intention initiale d'estimation plancher.

La méthode proposée est de consulter les experts (cellule de gestion de la CPFL) pour déterminer le nombre maximal de passages le plus probable, pour chaque site équipé d'éco-compteurs. En attendant l'équipe de recherche peut faire une estimation plancher non optimale. On note que notre problématique est très différente de celle des acteurs qui ont posé les éco-compteurs pour surveiller la fréquentation afin de prévenir les problèmes dus à une sur-fréquentation éventuelle : leur stratégie peut être alors d'avoir une estimation plafond raisonnable pour parer au scénario du pire.

Ainsi l'hypothèse est que si l'on divise par deux la somme de tous les passages, on ne dépassera pas le nombre réel de visites, ce qui correspond à une « estimation-plancher ».

(b) Corrections de données d'éco-compteurs

Les éco-compteurs ont plusieurs types de fragilités : aléas climatiques (brouillard, givre...), dégradations malveillantes ou non, pannes éventuelles. Les agents en charge de leur suivi ne peuvent réagir instantanément. Cela se traduit par des séries de données manquantes ou manifestement très sous-estimées.

Une correction de ces données a été effectuée par imputation sur la base de l'année précédente. Lorsque les données de l'année n sont très faibles dans l'absolu plus de x jours de suite (pour éviter l'explication par la météo), les données de l'année n-1 sont mobilisées.

Un taux de modulation est calculé sur l'ensemble de l'année entre années n et n-1, hors périodes défectueuses, pour tenir compte d'autres causes d'évolution : travaux sur le site, interdiction nouvelle de parcours etc. Ainsi les données de l'année n-1 pour la période suspecte ont été affectées du taux de variation calculé sur l'année (voir ci-dessous Pariou 2 : -19%).

Globalement ces corrections produisent une augmentation des chiffres 2019 de + 9%.

Le tableau complet des données d'éco-compteurs sur deux ans, et des corrections, est trop volumineux pour être annexé à ce document, mais il est intégré aux feuilles de la base de données. Un extrait de la feuille de calcul des imputations est donné ici :

| Eco-Compteur Pariou 2 (2018) | Pariou 2 (2019) avant imputation | Pariou 2 (2019) corrigé | Eco-Compteur Puy de Côme (2018) | Puy de Côme (2019) | Eco-Compteur Goules Sarcoui (2018) | Goules Sarcoui (2019) avant imputation | Goules Sarcoui (2019) corrigé |
|------------------------------|---|-------------------------|---------------------------------|--------------------|------------------------------------|--|-------------------------------|
| 51 914 | 42 038 | | | | 23 296 | 22 461 | |
| coeff correction | mais autres doutes (Mai ?) coeff sévère | -19% | | | | | 0% |
| 15 | 78 | 78 | 0 | 0 | 0 | 0 | 0 |
| 46 | 47 | 47 | 0 | 0 | 0 | 0 | 0 |
| 8 | 123 | 123 | 0 | 0 | 0 | 2 | 2 |
| 17 | 98 | 98 | 0 | 0 | 0 | 0 | 0 |
| 20 | 73 | 73 | 0 | 0 | 0 | 0 | 0 |
| 32 | 82 | 82 | 0 | 0 | 0 | 0 | 0 |
| 165 | 14 | 14 | 0 | 0 | 0 | 0 | 0 |
| 12 | 16 | 16 | 0 | 0 | 0 | 0 | 0 |
| 25 | 7 | 7 | 0 | 0 | 0 | 6 | 6 |
| 21 | 11 | 11 | 0 | 0 | 56 | 3 | 3 |
| 16 | 36 | 36 | 0 | 0 | 25 | 2 | 2 |
| 0 | 62 | 62 | 0 | 0 | 6 | 6 | 6 |
| 0 | 96 | 96 | 0 | 0 | 80 | 1 | 1 |
| 0 | 18 | 18 | 2 | 1 | 166 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 8 | 15 | 15 |
| 3 | 0 | 2 | 0 | 0 | 5 | 19 | 19 |
| 26 | 0 | 21 | 216 | 0 | 9 | 6 | 6 |
| 6 | 0 | 5 | 0 | 0 | 22 | 0 | 0 |
| 24 | 0 | 19 | 8 | 0 | 14 | 42 | 42 |
| 62 | 0 | 50 | 0 | 0 | 9 | 30 | 30 |
| 32 | 0 | 26 | 9 | 0 | 31 | 3 | 3 |
| 5 | 0 | 4 | 1 | 0 | 11 | 7 | 7 |
| 31 | 0 | 25 | 1 | 0 | 29 | 0 | 29 |
| 90 | 0 | 73 | 4 | 0 | 48 | 0 | 48 |
| 22 | 0 | 18 | 9 | 0 | 16 | 0 | 16 |
| 6 | 0 | 5 | 0 | 0 | 11 | 0 | 11 |
| 93 | 0 | 75 | 15 | 0 | 79 | 0 | 79 |
| 496 | 0 | 402 | 122 | 0 | 285 | 0 | 285 |
| 13 | 0 | 11 | 8 | 0 | 11 | 0 | 11 |
| 14 | 0 | 11 | 4 | 0 | 18 | 0 | 18 |
| 61 | 0 | 49 | 2 | 0 | 37 | 0 | 37 |
| 1 361 | 761 | 1 558 | 401 | 1 | 976 | 142 | 676 |
| 39 | 0 | 32 | 9 | 0 | 34 | 0 | 34 |
| 24 | 0 | 19 | 5 | 0 | 7 | 0 | 7 |
| 160 | 0 | 130 | 7 | 0 | 0 | 0 | 0 |
| 247 | 0 | 200 | 18 | 0 | 0 | 0 | 0 |
| 13 | 0 | 11 | 2 | 0 | 0 | 0 | 0 |

Cette hypothèse de mode d'estimation-plancher, et ce travail de correction-par-imputation des données suspectes ont été soumis à l'avis des agents de la cellule de gestion du site, en charge, entre autres choses, du suivi des éco-compteurs¹.

¹ Réunion par visioconférence le 10 Février 2021 suivi de l'envoi d'un dossier pour validation.

Ce point pourra faire ultérieurement l'objet de séances de travail complémentaires pour affiner les corrections d'anomalies, pour affiner l'estimation du taux de détection pour chaque site, et traiter de la pérennité de l'outil, de sa stabilité dans le temps en tant que dispositif d'ensemble. Par exemple on peut proposer pour une méthodologie d'observation pérenne de s'appuyer sur une moyenne glissante des données sur plusieurs années.

(c) *Fréquentation (visites) par site d'enquête*

La sélection de sites¹ ou leurs points d'accès est basée sur leur localisation (doit être situé dans la CPFL) et sur leur part dans la fréquentation de la CPFL. On dispose des données mensuelles des éco-compteurs et celles du panoramique des dômes, de Lemptégy, de la Grotte de la pierre, de la Maison du PNR (Tableau 3). En ce qui concerne Vulcania et Volvic, on ne dispose que des données de la fréquentation annuelle. La ventilation mensuelle de cette fréquentation est faite comme ci :

- Il a été supposé que la courbe de fréquentation annuelle de Vulcania et de Volvic est la même que celle du Panoramique (une autre approche peut être que cette courbe suit celle du PDD ou de Lemptégy + panoramique ou de Lemptégy +PDD) sauf les mois de Janvier, Février et Décembre où Vulcania est fermée (Janvier, Février, Mars et Décembre pour Volvic).
- Dans un premier temps, les pourcentages de répartition mensuelle de la fréquentation du Panoramique correspondant aux mois d'ouverture de Vulcania ont été additionnés. Dans un second temps, à partir de cette somme, les pourcentages de répartition de chacun des 9 mois d'ouverture de Vulcania (8 mois d'ouverture dans le cas de Volvic) ont été calculés. Ces pourcentages sont appliqués aux données de la fréquentation annuelle pour avoir les données mensuelles.

Enfin le lac d'Aydat est un des points très fréquentés de la CPFL mais on ne dispose d'aucune donnée sur cette fréquentation. Ce site sera tout de même intégré à la liste des points d'enquête.

¹ Ici, site désigne le lieu sur lequel s'effectue une ou des activités touristiques (gratuites ou payantes)

Tableau 3 : Sources des données par site d'enquête

| Sites | Sources des données |
|---------------------------------|--|
| Puys de Dôme | Billetterie du Panoramique + éco-compteur sentiers muletiers + éco-compteur chemin des chèvres (PDD1) + Puy de Côme |
| Parking Puy de la Vache | Éco-compteur Puy de la Vache 1 et 2 + éco-compteur Lassolas + éco-compteur Vichatel Gabion |
| Parking des Goules | Éco-compteur Goules Sarcouy + Pariou 1 et 2 |
| Lemptégny | Billetterie de Lemptégny |
| Vulcania | Billetterie de Vulcania |
| Parking Beauregard (2 parkings) | Éco-compteur puy de Jumes + Éco-compteur puy de Coquille + Éco-compteur puy de Louchadière + et du Puy de la Nugère |
| Parking Puys de la Combegrasse | Éco-compteur Combegrasse SO2 et Est 1 |
| Volvic | Fréquentation de l'espace d'information Volvic |
| Lac Aydat | Néant (Pas de données) |

Tableau 4 : Nombre de visites par sites et par mois en 2019

| Mois | Parking des Goules | Parking Puy de la Vache | Parking Puy des Gouttes et Chopine | Parking beau REGARD | Parking Puys de la CombeGrasse | Puy Dôme | de Lemptegy | Vulcania | Maison du PNR | Volvic | Total |
|-----------|--------------------|-------------------------|------------------------------------|---------------------|--------------------------------|----------|-------------|----------|---------------|--------|-----------|
| Janvier | 2 857 | 1 166 | 230 | 286 | 426 | 11 403 | 168 | 0 | 0 | 0 | 16 535 |
| Février | 7 240 | 8 715 | 423 | 719 | 1 270 | 25 868 | 1 503 | 0 | 0 | 0 | 45 737 |
| Mars | 6 093 | 3 520 | 1 337 | 2 631 | 1 039 | 21 160 | 2 872 | 10 028 | 0 | 0 | 48 679 |
| Avril | 9 287 | 4 202 | 831 | 1 042 | 864 | 46 588 | 14 536 | 31 179 | 1 508 | 8 885 | 118 920 |
| Mai | 8 488 | 6 590 | 1 377 | 2 714 | 1 027 | 46 445 | 15 512 | 29 132 | 1 610 | 8 301 | 121 194 |
| Juin | 9 656 | 5 808 | 1 054 | 957 | 1 535 | 58 995 | 15 087 | 40 656 | 2 284 | 11 585 | 147 616 |
| Juillet | 15 947 | 7 168 | 942 | 896 | 3 295 | 90 001 | 20 853 | 62 949 | 4 844 | 17 938 | 224 832 |
| Août | 27 983 | 9 788 | 1 592 | 1 375 | 3 761 | 143 620 | 32 342 | 99 141 | 6 935 | 28 251 | 354 788 |
| Septembre | 9 160 | 4 073 | 898 | 819 | 1 793 | 47 618 | 9 985 | 30 144 | 1 813 | 8 590 | 114 890 |
| Octobre | 7 928 | 4 730 | 337 | 660 | 1 215 | 28 563 | 9 040 | 16 992 | 516 | 4 842 | 74 821 |
| Novembre | 3 665 | 1 606 | 397 | 292 | 708 | 11 036 | 675 | 5 642 | 0 | 1 608 | 25 627 |
| Décembre | 3 827 | 1 886 | 412 | 282 | 886 | 17 003 | 0 | 0 | 0 | 0 | 24 295 |
| Total | 112 129 | 59 249 | 9 828 | 12 671 | 17 816 | 548 296 | 122 573 | 325 862 | 19 510 | 90 000 | 1 317 933 |

(2) L'échantillonnage par période et par lieu

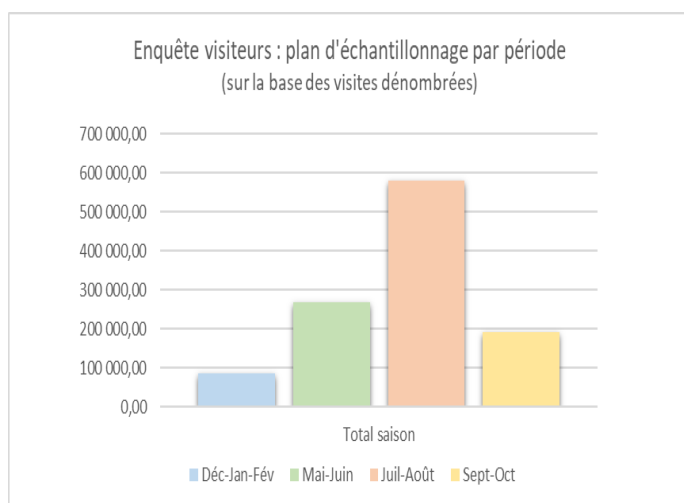
(a) Concernant les périodes d'enquêtes

Les périodes pré-identifiées étaient celles des mois de l'année, subdivisés selon les périodes touristiques particulières : vacances scolaires, WE et jours fériés, jours de semaine.

L'hypothèse est que cette saisonnalité affecte les caractéristiques des visiteurs et le comportement de consommation, elle constitue donc l'une des principales variables de stratification.

Comme indiqué précédemment, l'intensité d'enquête par période a été conçue pour refléter autant que possible la courbe des niveaux de fréquentation, sur la base de données de fréquentation mensuelle ou annuelle⁴³ de la CPFL fournie par des éco-compteurs⁴⁴ d'une part et par les sites disposant de moyens de comptage, par billetteries ou autres (le Panoramique des Dômes, Lemptégy, Vulcania, Volvic, Grotte de la pierre, maison du PNR), d'autre part.

Le plan d'échantillonnage par période était le suivant :



⁴³ L'entreprise Vulcania n'ayant communiqué que sa fréquentation annuelle, nous avons réparti ses chiffres sur les mois de l'année en suivant les proportions de la fréquentation du Panoramique des Dômes, en tenant compte des périodes d'ouverture respectives

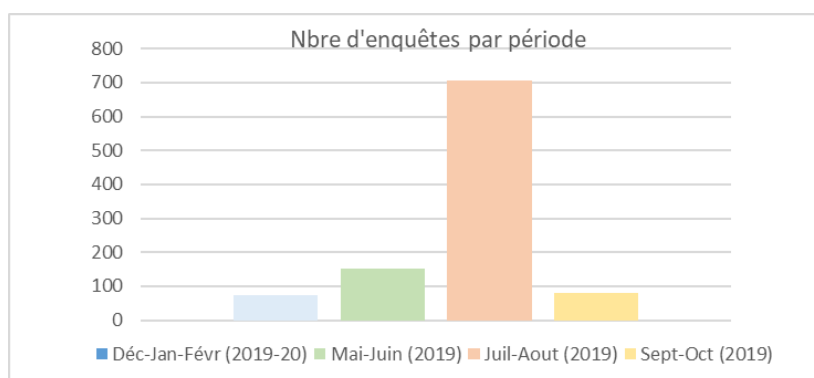
⁴⁴ Les éco-compteurs sont installés tout le long des chemins de randonnées menant vers les principaux puys de la CPFL.

Hormis le cas des jours fériés, le choix des jours de semaine a visé à couvrir tous les jours de la semaine mais à privilégier la fin de semaine en considérant qu'en cas de séjour hebdomadaire les réponses données en fin de séjour sont plus précises et plus fiables sur les activités pratiquées et sur les dépenses effectuées.

Périodes et jours d'administration de l'enquête :

Le nombre d'enquêtes par période s'établit ainsi

| Période | Nbre enquêtes |
|------------------------|---------------|
| Déc-Jan-Févr (2019-20) | 73 |
| Mai-Juin (2019) | 152 |
| Juil-Aout (2019) | 707 |
| Sept-Oct (2019) | 79 |
| total | 1011 |



Il apparaît un parallélisme dans les deux courbes (prévu/réalisé) mais aussi une certaine différence dans les proportions pour la période Juil-Aout notamment. Cette différence due aux contingences en termes de moyens d'enquête *est corrigée ultérieurement par la technique de redressement d'échantillon (voir section infra)*. Ainsi l'excédent d'enquêtes réalisées en haute saison ne biaisera pas les résultats, mais il permet d'avoir une meilleure « résolution » sur cette période importante.

Les 47 jours d'enquête se sont répartis de la façon suivante :

Par périodes :

- Hors vacances scolaires ou jours fériés
- Vacances scolaires
 - o 2019 : vacances d'été, d'Octobre, de Noël et fin d'année
 - o 2020 : vacances de Février
- Jours fériés :
 - o 2019 : Ascension, 14 Juillet, 15 Aout

N'ont pas été possibles, à cause de la COVID les enquêtes prévues en 2020 pour les vacances scolaires de printemps et les jours fériés de Mai (premier confinement).

En termes de répartition au sein des jours de semaine, le tableau est le suivant :

- Les Lundis : 4 jours d'enquête,
- les Mardis : 7 jours,
- les Mercredis : 8 jours,
- les Jeudis : 8 jours
- les Vendredis : 10 jours,
- les Samedis : 3 jours,
- les Dimanches : 7 jours

(b) Concernant les sites d'enquête

La sélection de sites⁴⁵ ou de leurs points d'accès est basée sur leur localisation (doivent être situés dans la CPFL) et sur leur poids de fréquentation.

Une diversité importante de sites a été recherchée⁴⁶, 15 lieux d'enquêtes ont été pratiqués.

Les 15 points d'enquête se répartissent ainsi (du Nord au Sud) :

- Parking Beauregard
- Volvic : parking et parc de la source près du pavillon des eaux de Volvic et parking de la Maison de la Pierre
- Parking des Gouttes
- Volcan de Lemptégy
- Vulcania

⁴⁵ Ici, site désigne le lieu sur lequel s'effectue une ou des activités touristiques (gratuites ou payantes)

⁴⁶ En tenant compte de contraintes pratiques (déplacements au sein d'une même journée) ou conjoncturelles (canicule extrême qui a limité les enquêtes sur parking pendant certaines périodes).

- Parking des Goules
- Puy de Dôme (4 points)
 - o Gare de départ du train Panoramique (montées, ou retours)
 - o Gare d'arrivée du train (retours)
 - o Sommet du Puy de Dôme (extérieurs ou intérieur de la maison du Sommet)
 - o Bas du sentier des Muletiers au col de Ceysat : (départs et retours)
- Maison du PNR (parking ou intérieur)
- Parking des Puys de la Vache et de Lassolas
- Parking du Puy de Combegrasse
- Plage et OT au lac d'Aydat

Là aussi l'intensité d'enquête par site a été conçue pour refléter autant que possible la courbe des niveaux de fréquentation, sur la base de données de fréquentation mensuelle ou annuelle de la CPFL fournie par des éco-compteurs⁴⁷ d'une part et les sites disposant de systèmes de comptage (le Panoramique des Dômes, Lemptégy, Vulcania, Volvic, la Grotte de la pierre, la maison du PNR) d'autre part.

Il faut signaler que Vulcania a « rationné » le nombre d'enquêtes autorisées sur son site (4 séances d'enquête au maximum et seulement en fin de journée). Même si des visiteurs de Vulcania peuvent avoir été rencontrés sur d'autres lieux d'enquête, ce qui est le cas pour toutes les activités, ce rationnement a créé une sous-représentation qui aurait nécessité un redressement excessif (coefficient 20 pour une des unités de redressement); pour cette raison principalement, il n'y a pas eu de redressement par lieu d'enquête et l'échantillon reste donc déséquilibré sur ce point.

b) Population ciblée et administration du questionnaire

Tous les types de visiteurs étaient ciblés a priori.

Deux questions se posent à ce sujet :

- *Que les choix de lieux d'enquêtes n'opèrent pas d'exclusion d'un type ou d'un autre.* Sur ce point les 15 lieux d'enquête sont un éventail large dans leur diversité. Des limites viennent du fait qu'ils ne couvraient pas les parkings sauvages. Par ailleurs les parkings de bus lorsque ceux-ci sont séparés du reste des véhicules (cas du parking du panoramique) ont été ignorés, ce qui a pu renforcer la sous-représentation des groupes dans l'échantillon.

⁴⁷ Les éco-compteurs sont installés tout le long des chemins de randonnée menant vers les (sommets des) principaux puys de la CPFL.

- *Que le processus d'enquête sur un lieu donné permette la rencontre aléatoire des visiteurs.*

La conception d'un mécanisme de sélection des visiteurs à enquêter est difficile dans la mesure où il n'y a pas de mécanisme de contrôle à l'entrée des sites. Les enquêteurs ont interrogé toute personne (seule ou faisant partie d'un groupe de visiteurs) qui se montre disponible et qui accepte de répondre à l'enquête. Ils ont l'instruction de ne pas privilégier un type particulier de visiteur, ni un secteur particulier du lieu d'enquête.

Les comportements très divers des visiteurs peuvent limiter les choix d'interlocuteur dans l'instant. Par exemple dans un lieu comme la gare de départ du Panoramique des Dômes les visiteurs allant vers les quais en urgence ne sont pas abordables, de même que, comme ailleurs, ceux qui vont vers les toilettes, ou qui s'occupent de leurs plus jeunes enfants. Par ailleurs certains membres de plus grands groupes (sorties organisés) étaient parfois plus difficiles à isoler du groupe.

Au final la présence simultanée de plusieurs enquêteurs (de 1 à 3, très majoritairement 2), de caractéristiques souvent différentes (genre et âge), ce qui permet de minimiser certains biais, et les instructions qui leur étaient données d'éviter autant que possible tout biais dans le choix des personnes à qui ils s'adressaient, permettait de rencontrer les visiteurs simultanément en plusieurs points du lieu d'enquête, et d'assurer autant que possible un aléa dans le tirage de l'échantillon enquêté.⁴⁸

Les visiteurs se présentent seuls ou par groupe. Un groupe de visiteurs est défini comme un ensemble d'individus qui visitent ensemble le site lieu d'enquête. L'objectif est d'interroger, pour un groupe donné, une personne en mesure de répondre pour l'ensemble du groupe ou pour un sous-groupe à préciser : notamment, grâce à une question ad hoc il était possible de distinguer au sein du groupe les personnes pour lesquelles le répondant interrogé donnait les informations budgétaires (souvent parce qu'elles partageaient les coûts du voyage, cas notamment du groupe familial). Deux variables étaient ainsi renseignées : « effectif du groupe » et « effectif concerné par le budget ».

⁴⁸ Enfin il est à noter qu'un choix est opéré dans le cas de groupes « mixtes » au sens de touristes accompagnés de parents ou d'amis locaux qui eux sont donc excursionnistes. Ne pouvant poser successivement les deux types de questionnaires au même groupe, l'enquêteur choisit l'une ou l'autre partie du groupe. Faute d'outil de suivi, il peut avoir privilégié par exemple le sous-groupe des touristes, ce qui constituerait un biais. Dans la base de données, au maximum 18 enquêtés ayant répondu être hébergés chez des parents ou amis, et dont les effectifs de groupe dépassent les effectifs budget, sont concernés, soit 1,8 %, sans savoir si l'autre partie du groupe est constituée d'autres touristes ou pas, donc une partie (seulement) aurait pu être adressée à des excursionnistes. D'autres cas de groupes mixtes sans hébergement commun, sont possibles.

Les enquêteurs étaient dotés d'un badge et se présentaient comme membres de l'Université, présentaient le projet et en particulier l'objectif économique de l'enquête. Il est à noter que très peu de visiteurs interrogés ont refusé d'aborder les questions sur leurs dépenses.

Les refus ont été notés de façon systématique, pendant certaines semaines d'enquête (13 semaines ayant renseigné 676 enquêtes sur les 22 semaines au total). A côté de ces 676 réponses positives, les enquêteurs ont noté 101 refus soient 13% du total, chiffre sans doute sous-estimé faute de prendre le temps de noter un refus et ses détails lorsqu'une opportunité se présente d'interroger quelqu'un d'autre. Bien que sous-estimé ce taux traduit le fait que les refus étaient minoritaires.

c) Effectifs enquêtés

Au total **1011 enquêtes** ont été effectuées.

L'effectif des groupes concernés est de :

3 474 visiteurs (avant redressement).

L'effectif pour lequel les dépenses ont été déclarées (dit « **effectifs budget** ») est de :

2 894 personnes (avant redressement) : 2 140 adultes et 754 enfants.

C. Saisie et premiers traitements des données

Cette section passe en revue les opérations de saisie et de premier traitement des données. Les principes sont exposés et des illustrations sont données. *Plus de détails sont donnés en annexe.*

1. Saisie sur le terrain

Les enquêteurs étaient munis de tablettes sur lesquelles était chargé le questionnaire sous le logiciel Limesurvey.

Le logiciel de saisie présente les questions selon un format imposé, assorties des formats et des règles que l'équipe a programmées :

- nature de la donnée attendue (numérique, choix fermés pré-codés ou saisie libre, etc.)
- bifurcations selon la réponse à une question d'aiguillage (par exemple « être-vous touriste ? » (ou excursionniste) envoyant vers deux séries de questions différentes.

Chaque session d'enquête voit ainsi se constituer un fichier de réponses dans l'ordinateur que les enquêteurs ont emporté sur place.

2. Déchargement des données, ajout de données de traçage

Juste après chaque enquête, ces fichiers sont déchargés en format Excel et des données « de traçage » sont ajoutées pour faciliter le suivi de l'enquête (N°indiv, Date_Enquete, Semaine, Mois, Questionnaire LimeSurvey). Deux de ces données sont à commenter :

- N°indiv résulte d'une numérotation finale synthétique, nécessaire car Limesurvey attribue de façon autonome un numéro d'ordre pour chaque ordinateur utilisé localement sur les lieux d'enquête
- « Questionnaire LimeSurvey » est utile pour faciliter la mise relation des résultats avec la version du questionnaire utilisée, sachant que pendant Mai et Juin 2019 (phase de mise au point) ces version ont évolué.

Ainsi une base de données est constituée, sous format MS Excel.

3. Statut des données

Pour chaque question les réponses peuvent être absentes car sans objet, ou parce que la version du questionnaire de la période ne comportait pas la question, ou parce que l'enquêteur a omis la question (oubli ou erreur de manipulation de la tablette) ou parce que l'enquêté ne savait pas ou ne voulait pas répondre.

Un codage de ces blancs dans toute la base de donnée a été réalisé sous la forme de trois codes : SO (sans objet), QA (question absente), N/A (réponse absente pour autres motifs).

4. conversions de données

Certaines données appellent une ou plusieurs conversions avant traitement.

Intervalles de dépenses

Pour chaque poste de dépense, les réponses sont saisies en cours d'enquête sous deux formes alternatives : soit par le choix d'un intervalle de valeurs, soit par la saisie d'un nombre. Le format « intervalles » était présent pour faciliter les réponses en cas d'hésitation. Mais pour effectuer les calculs ultérieurs, la conversion des intervalles en valeurs numériques est nécessaire. Pour chaque intervalle la valeur centrale a été choisie : par exemple « entre 50 et 100 euros » donne la valeur « 50 euros ». A noter que lorsque l'enquêté est gêné par l'intervalle car il pense être proche d'une des limites, l'enquêteur a pu passer alors en mode de saisie de valeur numérique.

Les intervalles de montants de dépenses comportaient toujours un dernier choix « plus de ». Lors de la conversion après enquête, il a été ajouté une dernière tranche supérieure dont la valeur centrale a traduit les réponses « plus de ... »⁴⁹. A elle seule cette conversion constitue un écrêtement éventuel des valeurs inconnues (car elles peuvent être encore supérieures). On peut la combiner avec une imputation lorsque c'est possible (voir infra). En tout état de cause, les réponses effectuées directement par l'enquêté en valeur numérique échappent à cette conversion.

Effectifs

Quatre questions portaient sur les effectifs à considérer :

- Etes-vous venu accompagné pour votre visite / séjour ?
- Les budgets indiqués ci-dessous sont-ils pour l'ensemble de votre famille / groupe / couple ou seulement pour vous ?
- A combien de personnes s'applique ce budget (y compris vous)?
- Combien d'adultes sont avec vous (y compris vous)?
- Combien d'enfants sont avec vous ?

La conversion a constitué à renseigner, à partir des réponses, le tableau suivant :

⁴⁹ Par exemple les dépenses d'hébergement déclarées sous forme de tranches ont pris la forme « plus de 900 euros » dans 24 cas sur 719 réponses de touristes, soient 3%. Le chiffrage converti était « 1 000 » (euros).

| Effectif du groupe concerné par le budget | Effectif groupe <i>adultes</i> concerné par budget | Effectif groupe <i>enfants</i> concerné par budget | Effectif du groupe (global) |
|---|--|--|-----------------------------|
|---|--|--|-----------------------------|

Sur 1011 réponses, 16 cas ambigus ont nécessité d'arbitrer sur les valeurs à affecter.

Durées de séjour

D'autres données ont fait l'objet de conversions, notamment celles relatives à la durée de séjour :

- la question était posée en nombre de nuitées. Pour certains calculs il était nécessaire d'utiliser des nombres de jours. Dans un premier temps la valeur retenue, paramétrée, est Nbre jours = Nbre nuits +1.
Mais grâce aux informations recueillies en 2021 par une enquête en ligne, une simulation plus fine a été faite (voir infra au chapitre « imputation »)
- Les lieux d'hébergement étant demandés par ailleurs, les durées de séjours ont été ventilées par département.

5. Vérifications / corrections

Des corrections de données saisies sont nécessaires. En effet la saisie peut être erronée soit par *faute de frappe*, soit par saisie dans une *mauvaise case* : notamment les dépenses, pour chaque type de dépense, sont renseignées successivement pour le jour de l'enquête puis pour le séjour : il est possible de confondre les cases voire de permuter les saisies (lorsque la discussion occasionne des retours en arrière). *Enfin des erreurs peuvent être le fait de l'enquête lui-même.*

Les corrections sont effectuées soit de façon systématique (par examen de valeurs extrêmes, ou selon des formules de test), soit au fil des opérations de conversion et d'imputation (cf infra).

Les tests systématiques ont porté notamment sur :

- la comparaison des dépenses déclarées par l'enquête pour le séjour par rapport à celles qu'il a déclarées pour le jour de l'enquête : on repère les cas d'inversion, ou d'oubli de renseigner la valeur pour le séjour alors qu'une valeur non-nulle est donnée pour le jour, etc.
- les effectifs, les durées, les nombres de venues dans la CPFL, qui ne peuvent être nuls.
- Les effectifs qui doivent être >1 si le groupe déclaré est différent de « personne seule », les effectifs du groupe qui doivent être > aux effectifs concernés par le budget, etc.

- pour les touristes itinérants, les durées de séjour en Auvergne, qui ne peuvent être inférieures à celles déclarées dans le PDD.
- les réponses concernant le poids de la CPFL dans le choix de destination, comparées aux réponses à la question « *Avez-vous connu le site CP avant votre séjour ou une fois sur place ?* ».
- les activités pratiquées, et certaines absences détectées selon le lieu d'enquête (p. exemple un enquêté au sommet du Puy de Dôme devrait avoir déclaré un accès à pied ou par le Panoramique).
- Etc.

Les vérifications *ponctuelles* recourent à une vérification « en ligne » par laquelle les autres informations sur le même groupe enquêté sont mobilisées :

- soit des commentaires ou des détails lorsqu'ils sont prévus dans le questionnaire (*cas des questions sur les activités, et sur chaque poste de dépense*),
- soit des données d'effectif, ou de durées, ou autres, qui peuvent contredire la valeur mal saisie.
- Par exemple un cas vérifié sur le mode de tourisme (itinérant ou non ?) est tranché grâce au fait qu'il est hébergé « chez de la famille ou amis ».
- A noter que les réponses de quelques enquêtés (déclarés touristes mais sans nuit passée, ou avec lieu d'hébergement trop éloigné pour une excursion à la journée), ont entraîné la création d'un type de visiteur non prévu au départ : les touristes « *en étape* », c'est-à-dire en transit du domicile vers leur lieu de vacances (éloigné de l'Auvergne) ou d'un lieu de vacances à un autre lieu de vacances, et qui font étape dans la CPFL, pour une petite demi-journée.
- Enfin un champ de la base de données en fin de questionnaire recueille les commentaires de l'enquêteur sur l'entretien et éventuellement sur le groupe enquêté : lorsque ce champ est suffisamment renseigné il peut permettre de mieux interpréter les données atypiques et/ou de mieux détecter les erreurs.

Ces divers moyens de recoupement sont utiles notamment pour juger des valeurs extrêmes, on citera par exemple :

- valeurs extrêmes (par personne et par nuit) d'hébergement en hôtel : le recoupement avec le lieu d'hébergement permet de vérifier la présence d'établissements, l'un économique, l'autre haut de gamme
- valeur extrême de repas pris au restaurant : le commentaire pris en note confirme que les enquêtés avaient l'intention de faire un repas « *exceptionnel* »
- valeur haute d'achat de produits locaux : il s'agissait de professionnels de la restauration qui ont acheté un lot de couteaux à Thiers
- Etc.

Lorsque la valeur de correction n'est pas évidente, on procède à ce qui est de l'ordre de « l'imputation » : donner une valeur jugée pertinente, probable, meilleure que la valeur qu'elle remplace. *L'explication générale en a été donnée ci-dessus en section II.3.2.1.* C'est l'objet de la section qui suit.

6. Imputations

Lorsqu'elle est possible, l'imputation vise à renseigner une valeur manquante, ou à corriger une valeur erronée. La section II.3.2.1 présente les méthodes rencontrées dans la littérature. Ici diverses méthodes ont été utilisées selon la nature des données. Il s'agit le plus souvent de prendre, pour remplacer une valeur manquante, une valeur moyenne calculée sur les autres enquêtes. Mais *des segmentations préalables ou des calculs préalables* sont effectués ; concernant :

- les dépenses d'hébergement :
 - les valeurs moyennes (par personne et par nuit) sont calculées *pour chaque type d'hébergement*, et leur produit par les effectifs et la durée est appliqué en remplacement de valeur absente (N/A)
- les dépenses groupées (appelées pension complète, demi-pension, package, forfait) :
 - une décomposition est opérée sur la base de dépenses estimées de restaurant (hypothèse) ou de dépense d'activité recalculée, voir infra (par exemple « pack hotel+Vulcania »)
- les dépenses d'alimentation :
 - Une estimation préalable des nombres de repas pris au restaurant est effectuée.
 - Une estimation de prix moyen de repas est effectuée sur les dépenses de restauration *du jour* divisées par 1,5 (entre 1 et 2 repas pris en compte), puis cette estimation est appliquée pour estimer le nombre de repas pris *pendant le séjour* en divisant le budget séjour par le coût moyen calculé sur les jours.
 - Puis une valeur moyenne d'alimentation par personne et par jour est appliquée sur le nombre de repas liés à la durée du séjour (nombre théorique diminué du nombre estimé de repas pris au restaurant)..
- les dépenses d'activités de loisir :
 - pour toute activité payante déclarée (sites payants de la CPFL, activités sportives, nautiques, ânes, montgolfière, musées, zoo, etc.), *les tarifs des établissements concernés* (dans la CPFL ou ailleurs) ont été recherchés et les calculs effectués (tenant compte de la différence entre adultes et enfants).

- les dépenses de transport (non compris les trajets des touristes pour aller de leur résidence permanente à leur lieu d'hébergement) :
 - o Pour les excursionnistes un calcul est fait à partir du lieu de résidence des excursionnistes pour aller vers un point de la CPFL (par convention la commune d'Orcines). L'application Via Michelin fournit l'estimation essence + péages (éventuels).
 - o Dans le cas des touristes le calcul a été fait également puisque l'on connaît leur lieu d'hébergement. Cependant comme l'essence des touristes peut avoir été achetée hors d'Auvergne la non-réponse a été transformée en valeur nulle ce qui évite de surestimer l'estimation des dépenses locales. L'imputation possible a été calculée, elle aurait augmenté l'estimation de l'impact économique total de +0,54 %).

7. Simulations de compléments ou raffinements éventuels

a) Cas de la durée de séjour en nombre de jours

Comme mentionnée ci-dessus en section « conversion de données », l'enquête comportait une question sur les durées de séjour exprimées en nuitées. Les calculs ont été effectués en tenant compte lorsque nécessaire d'une durée en jours estimée à nombre de nuitées+1.

En 2021 dans le cadre des travaux sur la conception de dispositifs d'observation reproductibles donc allégés, un test d'enquête en ligne a été effectué, et le questionnaire a inclus les deux questions en nuitées et en jour. Sur la base de 693 réponses, on constate des cas où nb jours = nb nuits,, des cas majoritaires où nb jours = nb nuits +1 (et des cas très minoritaires où nb jours = nb de nuits-1, cas qui seront négligés) . Il ne s'agit pas d'ajouter à tous les séjours une durée fractionnaire. Donc le choix a été de tirer aléatoirement une proportion d'enquêtés auxquels un jour est ajouté ou retranché. De plus une segmentation sommaire a été effectuée pour être plus proche de la réalité. Ainsi les taux sont de:

- Pour les courts séjours (<= 3 nuitées) : 75% de cas à nb jours = nb nuitées +1
- Pour les séjours > 3 nuits : 66% de cas à nb jours = nb nuitées +1

Le premier constat est que l'option initiale est largement majoritaire. Pour aller plus loin faut-il modifier les données par imputation aléatoire ? Cela biaiserait les analyses fines qui feraient intervenir les durées en jours croisées avec d'autres critères, donnant des sous-populations peu nombreuses sensibles à l'approximation aléatoire. Le compromis retenu est d'exploiter la simulation dans certaines interrogations (durées moyennes de séjours) et de laisser la BDD en état pour le reste.

b) Cas des « touristes répétés »

En section IV.2.5.1 infra, sera traité le cas des touristes qui viennent plusieurs fois dans l'année en Auvergne, information non connue. Un taux issu du SDT 63 pour l'année 2018 permet de faire une simulation de leur prise en compte. Mais comme cette simulation repose sur l'attribution aléatoire de visites supplémentaires à un dixième des touristes, la BDD ne sera pas modifiée mais les résultats de la simulation sont pris en compte dans l'appréciation des intervalles de confiance.

c) Cas des « Effectifs-budget » versus « Unités de Consommation »

Les imputations recourent à des valeurs moyennes par personne, ou par personne et par nuit. Parfois ces moyennes font l'objet de typologies ou de calculs préliminaires (types d'hébergement, calcul du nombre de repas pris au « restaurant »). Pour calculer des moyennes de dépense par personne (et par nuitée ou par jour), les « effectifs-budget » sont pris en valeur totale (adultes + enfants). *Dans les calculs présentés, un groupe d'une composition donnée se verra donc affecter une valeur moyenne calculée sur des groupes de toute composition, excepté dans le cas des activités payantes pour lesquelles les tarifs différenciés adultes/enfants ont été pris en compte. C'est évidemment une simplification, connaissant tous les travaux qui s'attachent à rendre compte au mieux des revenus ou des consommations en tenant compte de la composition des ménages/familles ou autres groupes de consommation. Ces travaux traitent « d'Unités de Consommation (U.C.) » pour neutraliser les différences entre adultes (selon leur rang), entre enfants (selon leur âge) et entre adultes et enfants.*

Plusieurs échelles ont été proposées pour ramener à un équivalent UC les divers membres du groupe (familial en général) : premier adulte, autres adultes, enfant (selon son âge ou non).

- l'échelle d'Oxford est la plus connue bien que déjà ancienne. Elle propose :
1^{er} adulte = 1 autre adulte = 0,7 enfant = 0,5
- l'échelle dite modifiée propose :
1^{er} adulte = 1 2^e adulte ou enfant = 0,35
- l'échelle OCDE de « niveau de vie », utilisée par le Credoc notamment :
1^{er} adulte = 1 autre adulte ou enfant de plus de 14 ans = 0,5 enfant < 14 ans = 0,3

Ces échelles présentent toutes l'inconvénient de ne pas traiter spécifiquement de consommation touristique. Il est évident que le budget annuel d'un ménage diffère de la seule consommation de ses vacances. Mais le principe d'une dégressivité au sein du groupe (familial ou autre) est à tester.

Une simulation a été effectuée en adoptant l'échelle d'oxford.

Le total des UC non redressées sur l'échantillon (2 178) est bien différent du total des effectifs (2 894), come attendu. A noter que ce total se rapproche du total des adultes utilisé par le SDT pour calculer des dépenses moyennes par adulte et par nuit (2 140).

La simulation s'applique d'abord aux imputations basées sur des valeurs moyennes (hébergement, restauration, alimentation). Les écarts sont très faibles :

| Hébg | Restau. | Alim. |
|---------|---------|--------|
| 220 233 | 109 519 | 58 958 |
| 220 696 | 109 927 | 58 846 |
| 0,2% | 0,4% | -0,2% |

La simulation sera évoquée ci-après au stade des calculs d'effets économiques globaux (test de possibles différences quant aux UC moyennes vs effectifs moyens, entre les 4 populations distinguées dans ces calculs : touristes/excursionnistes et locaux/non locaux).

Enfin elle demeurera pertinente pour tester des interrogations particulières plus fines au sujet des dépenses.

8. Comparaisons déclaré/imputé, corrections

Sur certaines données, les calculs de type « imputation » ont été effectués systématiquement :

- Dépenses d'activité
- Dépenses de transport

Au-delà du travail d'imputation pour les réponses absentes (N/A), les résultats comparés aux dépenses déclarées ont entraîné ou pourraient entraîner des arbitrages donc des corrections de type systématique:

- en ce qui concerne les activités, le max (déclaré, calculé) a été retenu : si l'imputation est supérieure on considère que c'est le signe d'une erreur de chiffrage par l'enquête, si l'imputation est inférieure on considère que c'est le signe d'oublis dans la liste des activités pratiquées
- en ce qui concerne les dépenses de transport on pourrait également faire un test systématique et prendre le Max (déclaré, imputé) mais l'essence peut avoir été achetée hors d'Auvergne, aussi il vaut mieux s'en tenir aux déclarations des enquêtés (la question portait sur les dépenses locales)

Plus généralement les travaux d'imputation ont aussi occasionné des vérifications et des corrections ponctuelles au cas par cas.

9. Traçabilité des opérations

La base de données garde la trace des conversions, corrections, imputations par plusieurs dispositifs :

- Une première feuille du fichier Excel est intitulée « toutes réponses initiales », la seconde feuille intitulée « feuille corrigée » comporte les valeurs corrigées le cas échéant.
- Dans cette seconde feuille certaines variables font l'objet de deux colonnes : « réponses initiales » et « après corrections »
- Dans la plupart des feuilles les valeurs corrigées sont notées en couleur rouge et la cellule correspondante peut porter un commentaire explicite sur la correction effectuée.
- Les feuilles de calcul « thématiques » (dédiée aux calculs sur un type de dépense notamment) comportent :
 - En lecture par colonne (variables utilisées), des lignes sous l'intitulé de colonne, lignes qui précisent les définitions, les particularités du calcul opéré dans cette colonne
 - En lecture par ligne (cas enquêtés), un champ destiné aux commentaires éventuels concernant le traitement de chaque enquêté

D. Redressement des données collectées

S'agissant d'une enquête par sondage et non d'un recensement exhaustif, il s'agit de tendre vers la meilleure représentativité de l'échantillon enquêté. En section III.1.3.1 on a présenté l'alignement du plan d'enquête (nombre, dates et lieux des enquêtes) sur les données connues en termes de visites de sites en CPFL (dates et lieux). A présent il s'agit d'améliorer la représentativité de l'échantillon réellement enquêté.

Globalement plusieurs méthodes sont possibles, dont la réduction des parties excédentaires dans l'échantillon (par suppression de réponses), ou la pondération des réponses, soit à l'échelle individuelle, soit par strates. La première méthode a l'inconvénient de diminuer l'information disponible, c'est la seconde méthode qui a été retenue.

1. Pondération probabiliste ex post

En section II.3.1.3 le problème d'échantillonnage lors d'enquêtes en milieu ouvert est exposé et la recommandation de calcul probabiliste (Eurostat, 2000) est signalée, avec partage des poids à l'échelle individuelle (Lavallée, 1995 ; Ardilly et Le Blanc, 2001 ; Leroy, 2018). Plus précisément (Ardilly, P., 2006 ; Maurence, E., 2012) ciblent les enquêtes dans lesquelles le *facteur temps* intervient : le visiteur qui revient x fois sur un lieu d'enquête a une probabilité $p(x)$ (fonction croissante) d'être enquêté, et doit être affecté d'un poids minoré en relation avec $p(x)$. Ardilly prend l'exemple d'une enquête en hôtel, où le client qui reste x jours doit être pondéré en proportion de $1/x$ ⁵⁰. On remarquera que ces pondérations ne peuvent être faites avant enquête (au stade de la constitution de l'échantillon à enquêter) faute d'information sur les pratiques de chaque visiteur.

Un premier redressement de l'échantillon, sous la forme d'une pondération individuelle probabiliste a ainsi été opéré, sur la base du nombre de visites de chaque groupe enquêté sur un lieu enquêtable.

Les effectifs de chaque groupe sont ainsi multipliés par : **1/nbre de visites du groupe**.

En alternative, il serait possible d'envisager le nombre de *jours* de visite de la CPFL, si l'enquête se limitait à un lieu par jour d'enquête, bien que le visiteur qui fréquente deux lieux le même jour ait tout de même deux fois plus de chances d'être enquêté.

Cependant cette pondération alternative peut être très utile si on veut procéder plus tard à des enquêtes simplifiées, car elle ne nécessite qu'une question sur « le nombre de jours passés en CPFL » et non pas le détail des visites sur site dénombrable.

⁵⁰ (Maurence, 2012) prend l'exemple d'une enquête dans le cadre d'un festival avec plusieurs événements sur plusieurs jours

Le calcul a été effectué, la corrélation entre les deux coefficients probabilistes est 0,78, l'estimation de la fréquentation serait inférieure (790 000 au lieu de 840 000 soit - 6%), mais celle des résultats économiques serait tout de même supérieure : +5% pour l'impact économique local, ce qui traduit des changements plus sensibles dans les variables intermédiaires.

2. Post-stratification par période ou par lieu-période

Dans le cadre de l'étude sur les dépenses liées sur la fréquentation de la CPFL, on ne dispose pas d'information concernant la population globale (des visiteurs de la CPFL tous sites confondus) sur les types de variables comme : les types de visiteurs (touriste ou excursionniste), l'effectif du groupe de voyage, les âges, le type d'hébergement, la durée du séjour, etc. En revanche, on peut avoir des informations sur la fréquentation mensuelle de la CPFL.

En faisant l'hypothèse que les dépenses varient selon le mois de fréquentation, notamment en faisant l'hypothèse amont que les *types de visiteurs* (familles, couples âgés, jeunes) sont différents selon les périodes, que leurs *types de dépenses* sont différents, ainsi que les *tarifs* de certaines consommations (hébergement, activités), on veut vérifier la représentativité de l'échantillon selon les périodes et le cas échéant le redresser selon ces périodes. On peut utiliser la proportion de la fréquentation mois par mois ou *période par période* pour le faire.

On peut aussi faire l'hypothèse que les dépenses varient en fonction des types de *pratiques* des visiteurs, pour lesquelles la seule information globale avant enquête est la fréquentation des *lieux ou sites de visite*. Mais ce redressement est plus complexe du fait qu'un même enquêté visite plusieurs lieux, cf infra.

C'est un chemin méthodologique comparable à celui qui avait été suivi pour construire l'échantillon d'enquête (échantillonnage « lieux-moments » cf section II.3.1.3). Mais *la différence entre l'échantillonnage avant enquête et le redressement après enquête* tient au fait qu'avant enquête on ne peut « aligner » qu'une courbe en nombre *d'enquêtes* par période et par lieu, sur la courbe connue des *visites* dénombrables en CPFL, par période et par lieu. Or à présent chaque entretien d'enquête « porte » un nombre *de visites* variable (nombre de sites dénombrables visités par le groupe considéré, multiplié par l'effectif du groupe), le plus souvent supérieur à 1. Après dépouillement de l'enquête, la phase de redressement, elle, tient donc compte des nombres de *visites* de chaque groupe enquêté : il y a cohérence sur les variables comparées.

Cette méthode de stratification a posteriori est donc possible puisqu'on dispose du tableau indiquant le nombre de visites chaque mois de chaque site d'enquête (cf tableau IV à la section II.2.3.1.1.). En regroupant les mois et les lieux en classes, on construit un nombre limité de croisements de classes, appelés URE (Unités de redressement). Ainsi, le rapport entre le poids de l'effectif théorique d'une URE dans le total, divisé par le poids au sein de l'échantillon de l'effectif enquêté de cette URE, constitue alors le coefficient multiplicateur à appliquer aux individus correspondants à l'URE.

Le redressement consiste à affecter des poids aux réponses des enquêtés dans l'analyse des résultats de telle sorte que l'on retrouve alors les proportions connues sur la population. Ce sont les poids de sondage des individus de l'échantillon (et, par prolongement, les estimateurs) qui

sont redressés, et non pas l'échantillon, dans le sens où toutes les réponses de l'échantillon sont conservées.

a) Redressement par période

Soit P_1 et p_1 la proportion de visites sur la CPFL en Hors-saison (Avril, Mai, Juin, Septembre, Octobre, Novembre) respectivement dans population totale et dans l'échantillon d'enquête ; P_2 et p_2 la proportion de visites sur la CPFL en Haute-saison (Juillet, Août) et P_3 et p_3 la proportion de visites sur la CPFL en Hiver (Décembre, Janvier, Février, Mars). Ces proportions p_1, p_2, p_3 de visites par période au sein de l'échantillon ont été calculées après application du redressement probabiliste individuel.

A partir d'une stratification a posteriori ou post-stratification sur le critère saison ou période on peut donner un poids aux questionnaires selon le type de période à partir de la formule :

Moyenne-saison

Avril, Mai, Juin, Septembre, Octobre,

$$Q_1 = P_1 / p_1$$

Haute-saison

Juillet, Août

$$Q_2 = P_2 / p_2$$

Hiver

Novembre-Décembre, Janvier, Février, Mars

$$Q_3 = P_3 / p_3$$

Il en résulte les coefficients de redressement stratifiés par période (qui seront ensuite multipliés par les coefficients individuels probabilistes mentionnés précédemment)

| scénario 5 | fréquentation (écocompteurs et billetteries) | | visites des enquêtés | | coeff. redr. / période |
|-----------------------|---|-------------|----------------------|-------------|---------------------------|
| | | | | | |
| Avr-Mai-Juin-Sept-Oct | 577 441 | 44% | 570 | 21% | 2,12 |
| Juil-Août | 579 619 | 44% | 1 911 | 69% | 0,64 |
| Nov-Déc-Jan-Fév-Mars | 160 872 | 12% | 282 | 10% | 1,20 |
| Total | 1 317 933 | 100% | 2 764 | 100% | |

Le test du Khi-deux a confirmé la nécessité du redressement.

NB : pour les excursionnistes répétés, les nombres d'excursions hors jour d'enquête (et donc de visites, selon l'hypothèse de 50% de visites sur lieux d'enquête, mentionnée infra) ont été ventilés par période au prorata des durées des périodes.

b) Redressement par lieu-période : abandonné

Le plan d'échantillonnage avant enquête reposait sur une distribution des visites selon deux dimensions : celle des périodes et celle des lieux.

Des travaux du même type que ci-dessus ont été effectués pour obtenir des coefficients de redressement selon un nombre donné de classes (appelées URE ou Unités de redressement) croisant des classes par période et des classes par lieux.

Il convient de ne pas multiplier les URE pour conserver au sein d'une URE le plus de chance de refléter la diversité réelle représentée par l'échantillon.

De plus il convient d'éviter des coefficients de redressement trop élevés qui font courir le risque de survaloriser des cas particuliers au sein d'une classe, surtout si celle-ci est de faible effectif.

Un essai a été effectué avec 9 classes croisant : 1- les trois périodes exposées dans la section précédente et 2- trois regroupements de lieux de fréquentation dénombrable et d'enquête. Sur ce point l'hypothèse était faite d'une certaine homogénéité de comportements au sein des visiteurs rencontrés sur les parkings et comptés sur l'accès à un Puy, d'une homogénéité possible entre visiteurs de sites (Vulcania, Lemptégy, Volvic, etc.). Enfin le cas des visiteurs du Puy de Dôme semblait particulier et une classe de lieux leur était réservée, confondant les accès par sentier et par Panoramique.

Sans affinement des calculs (notamment sur les excursions répétées, cf infra), le tableau des coefficients était :

| | Coefficients de redressement | | |
|-----------------------|------------------------------|---|--------------------------|
| scénario 2bc | Puy de Dôme | Parking des Goules/Parking Puy de la Vache/Parking Puy des Gouttes et Chopine/Parking beauregard/Parking Puys de la Combegrasse/Parking de Fontfreyde | Vulcania/Lemptégy/Volvic |
| Avr-Mai-Juin-Sept-Oct | 1,3 | 2,2 | 22,7 |
| Juil-Août | 0,7 | 0,3 | 0,7 |
| Nov-Déc-Jan-Fév-Mars | 2,0 | 3,8 | 1,1 |

Il apparaissait un premier écueil, celui du coefficient très élevé, au moins en proportion des autres, de 22,7 pour l'URE « moyenne saison à Vulcania, Lemptégy, Volvic ». Cette anomalie est liée au fait que Vulcania a contingenté les enquêtes autorisées. Par ailleurs concernant les excursionnistes répétés, on ne dispose pas de l'information sur le détail des lieux visités en cours d'année.

Enfin un grand nombre de visiteurs se rendent sur plusieurs des lieux au cours d'un séjour, voire d'une excursion⁵¹. Il semble difficile de calculer des coefficients tenant compte de ces visites multiples (cf infra la section IV.2)

Pour ces raisons, il n'a pas été procédé à un redressement incluant la dimension « lieux ». Il demeure que le déséquilibre du plan d'enquête sur cette dimension réduit la précision des résultats, et surtout celle d'analyses par lieux (exemples ci-dessous).

c) Cas particulier de Vulcania

Le biais dû au rationnement des enquêtes sur le site de Vulcania, non corrigé par redressement par lieu serait sensible dans les interrogations fines de la BDD : un exemple parmi d'autres serait le calcul des taux de touristes et d'excursionnistes comparant les proportions selon le lieu d'enquête et selon le site visité. Même si des différences sont attendues (par définition des deux calculs) elle est particulièrement importante pour Vulcania (63/37 parmi les enquêtés sur site, contre 83/17 parmi les visiteurs du site).

d) Cas particulier du Lac d'Aydat

Le lac d'Aydat est un site pour lequel n'existe pas de décompte systématique des visiteurs. Malgré ce point, il est notoire qu'il connaît une fréquentation importante, et 50 enquêtes y ont été menées.

En 2019, faute de chiffre à ajouter au nombre total de visites recensées, le calcul des effectifs globaux de visiteurs (voir infra) ne prend pas en compte les visites à Aydat. Mais le coefficient de redressement probabiliste, lui, prend en compte ces visites puisque c'est un lieu d'enquêtes.

A noter que parmi les enquêtés au lac d'Aydat, la moitié étaient des excursionnistes, ce qui même après fort redressement probabiliste (excursionnistes répétés) donne une proportion élevée de 36% d'excursionnistes (supérieure de 6-8 points à la moyenne générale).

3. Qualités et limites de l'enquête et des traitements de données, précision des résultats

a) Principe

Sur le plan méthodologique l'enquête présente des qualités qui donnent une certaine robustesse « a priori » aux résultats :

⁵¹ L'un des arguments de la direction de Vulcania était que l'enquête allait de toutes façons rencontrer des clients de Vulcania ailleurs, notamment sur le site du Puy de Dôme.

- Taille de l'échantillon
- Efforts d'échantillonnage selon lieux et périodes d'une variété importante
- Richesse d'informations permettant des recoupements et des vérifications nombreuses

Par ailleurs elle présente des limites diverses même si ces limites sont chacune de portée limitée.

Une partie de ces limites ont fait l'objet de simulations et non de corrections qui seraient incorporée aux données finales de la base de données car par exemple certaines simulations attribuent de façon aléatoire une valeur manquante, ce qui permet d'approcher un résultat global amélioré, mais fait courir des risques d'erreur sur des interrogations fines (croisements de variables) : c'est le cas de la simulation sur les touristes répétés.

Plus généralement, les traitements font intervenir des hypothèses et des approximations en cascade.

Au final il est donc très difficile de calculer un « intervalle de confiance » des résultats qui sont affichés dans la suite de ce rapport.

En solution palliative il a été rappelé que chaque simulation prise isolément donne l'indication de résultats différents de seulement quelques %.

Pour rappeler visuellement la notion d'intervalle de confiance, on pourra afficher plusieurs valeurs d'un résultat selon les hypothèses ou les méthodes alternatives :

Par exemple le redressement probabiliste pratiqué est celui qui a le plus d'effet quantitatif, puisqu'il divise la « représentativité d'une enquête par le nombre de visites déclarées.

Un redressement probabiliste moins intense diviserait cette représentativité par le nombre de jours déclarés comme passés sur le site.

La simulation montre des différences notamment sur le premier résultat qui est le calcul de proportions entre touristes et excursionnistes (différence de 4 points, modérée), donc sur les résultats économiques.

Mais cette simulation qui a été effectuée systématiquement sur toutes les interrogations de la base montre des différences plus faibles sur des résultats plus détaillés (par exemple la ventilation des touristes, ou des excursionnistes, selon la composition du groupe : couple, famille, amis, etc.).

On montrera les deux résultats sur un même graphique, concernant les proportions touristes/excursionnistes, bien qu'il ne s'agisse pas de la représentation d'un « intervalle de confiance ». L'intérêt de cet affichage tient également au fait que dans des enquêtes ultérieures, simplifiées, on peut ne demander que ce nombre de jours passés sur le site, plutôt que la liste des visites pratiquées. Ce redressement est donc intéressant à comparer.

b) Détail des limites ou biais de l'enquête et correctifs

Limites et moyens de prise en compte de ces limites

- Une période a été sous-enquêtée (Covid 2020): la moyenne saison (c'est pris en compte par le **redressement** mais la diversité des cas est moins bonne qu'en haute saison)
- Les manifestations sportives n'ont pas été enquêtées
- Le taux de dénombrement des excursions répétées (combien de visites sur sites dénombrables ?) est inconnu (**taux imputé 50%**)
- On n'a pas interrogé les touristes sur d'éventuels retours dans l'année (mais il est possible qu'alors la causalité CPFL soit faible) : **simulations** faites sur la base du SDT 63 2018 (taux 11%) : imputé +1 à +3 visites dénombrables à 10% des touristes
- La conversion nombre de nuitées vers nombre de jours n'est pas déclarée, **simulation** : a été imputée sur la base de l'enquête 2021
- Les visiteurs n'ont pas indiqué le nombre de visites faites sur un même site (par exemple à Aydat)
- Limites du redressement sur conformité au plan d'échantillonnage concernant les lieux d'enquête : problème des URE à trop faible effectif (moy. saison X Vulcania) => pas de redressement sur les lieux de visite et d'enquête

Biais d'enquête

- Les groupes (voyages organisés) sont peut-être sous-enquêtés (comme dans d'autres études de ce type) => chercher auprès des sites visités et auprès d'opérateurs de voyage toutes données utiles (TC Domes aurait **estimé** à 10% la proportion des groupes dans les passagers du Panoramique)
- Lorsqu'un groupe mixait des résidents locaux et des personnes accueillies, donc touristes, les touristes ont pu être privilégiés => une possible sous-estimation des résidents locaux (excursionnistes), et des résidents secondaires
- Et sans oublier les « biais d'enquêteurs » inévitables, quoique **limités** par le nombre d'enquêteurs (5) et leur diversité de profils.

IV. CPFL : résultats de l'enquête, partie I : éléments nécessaires au calcul d'effet économique

Ici ne sont présentées que les parties de l'enquête nécessaires au calcul des effets économiques, c'est-à-dire essentiellement celles qui concernent

- Le taux de motivation des touristes pour la CPFL dans leur choix de destination
- Les données permettant une estimation-plancher de la fréquentation totale de la CPFL.
- Les dépenses de chacun

Les autres résultats descriptifs, sur les profils ou les pratiques des visiteurs, ainsi que diverses analyses croisées font l'objet d'une Partie II infra.

A. Les motivations pour la CPFL : calcul du coefficient de « l'effet CPFL »

On appellera « effet CPFL » la relation de cause à effet entre l'existence de la CPFL et les grandeurs observées. Dans cette partie de l'étude Valo-Sites, les grandeurs observées sont les dépenses des visiteurs.

1. Les motivations des excursionnistes : hypothèse effet CPFL 100%

Concernant les excursions l'hypothèse est celle d'un motif global unique : si l'excursion se passe dans la CPFL c'est globalement dû à l'existence de la CPFL (et de tous ses attributs indifféremment), soit un coefficient d'affectation de 100%.

2. Les motivations des touristes : calcul de « l'effet CPFL » (touristes)

Comme indiqué ci-dessus dans la revue de littérature, les méthodes de calcul des effets économiques imputables au site prennent en compte, pour les touristes, la totalité de leurs dépenses de séjour mais en proportion de l'effet du site sur leur décision de venir dans leur destination. Ce point de méthode est donc central pour le calcul.

Cette détermination a des conséquences importantes dans les résultats en termes de dépenses affectables puisque « l'effet CPFL dans le choix de destination » joue sur l'ensemble du séjour et de ses dépenses, au-delà du jour d'enquête. Dans le cas extrême si le séjour devait ne pas avoir lieu ou pas ici en l'absence de la CPFL, même les activités du séjour qui ne se sont pas passées dans la CPFL ou en lien avec celle-ci, n'auraient pas eu lieu.

Deux approches sont proposées selon les auteurs : « l'attribution » qui détermine un coefficient unique ou souvent graduel, selon l'intensité estimée de l'effet-site, et « la substitution » qui ne retient que les dépenses des visiteurs qui ne seraient pas venus dans leur destination en l'absence du site. Les deux approches ne sont pas comparables, la deuxième étant dite « prudente » ou

minimale : Carlsen et Wood (2004) appliquent les deux et sur chacun de leur cas d'étude les résultats varient d'un facteur 10 environ.

La méthode suivie ici (voir la section II.2.1.2 Part des dépenses à prendre en compte) est celle de l'attribution et non de la substitution. Il s'agit de déterminer pour chaque visiteur une position dans une échelle de 0 à 100% et de prendre en compte la proportion correspondante des dépenses totales.

3. Question principale

Le critère utilisé dans la littérature est celui des réponses à des questions directes sur les représentations et les motivations des visiteurs. Dans les exemples d'application la plus simple de ce type de méthode (Réseau des Grands Sites de France, 2008), une seule information vise à trouver le coefficient, notamment la réponse à une question directe du type : « Votre choix du lieu de séjour est-il dû : – uniquement – principalement – à parts égales – secondairement – pas du tout à l'existence de la CPFL ? »

Considérant que le visiteur répondra plus facilement à des questions de type dichotomique (choisir A ou B) avec un troisième choix de type « autre », qui renvoie sur une question suivante, la question principale est posée en deux temps :

- « diriez-vous que vous êtes venus (dans votre destination) : **uniquement** à cause de la CPFL, ou **pas du tout** (pour une autre raison) ou **en partie ... ?** »
- puis « **si** vous êtes venus **en partie** pour la CPFL, est-ce **principalement, à titre secondaire, ou à égalité** (avec d'autres raisons) ? ».

Les réponses permettaient ainsi d'attribuer à cet effet un **coefficient C1** prenant l'une des 5 valeurs :

1 - 0,75 - 0,5 - 0,25 ou 0.

Les réponses ont été les suivantes, en effectifs concernés (redressés) :

| Diriez-vous que vous êtes venu dans cette destination | | |
|--|--------------|-------------|
| | Eff. redr2 | % |
| Uniquement pour la CP | 395 | 32,8% |
| En partie pour la CP | 558 | 46,3% |
| Autr. raisons / découv. s place | 252 | 20,9% |
| Total explicites | 1 206 | 100% |

| Si vous êtes venu <i>en partie</i> pour la CPFL, est-ce | | |
|--|------------|-------------|
| | Eff. redr2 | % |
| principalement pour | 204 | 37,4% |
| A parts égales | 182 | 33,4% |
| secondairement pour | 158 | 29,1% |
| Total explicite | 544 | 100% |

Et au final en combinant les deux réponses :

| => sont venus dans cette destination ... | | |
|--|--------------|-------------|
| | Eff. redr2 | % |
| Uniquement pour la CPFL | 395 | 33,2% |
| principalement pour ... | 204 | 17,1% |
| A parts égales | 182 | 15,3% |
| secondairement pour ... | 158 | 13,3% |
| Pour d'autres raisons | 252 | 21,2% |
| Total | 1 192 | 100% |

4. Question de contrôle

Au vu de l'importance du rôle de ce coefficient, les réponses ont été contrôlées par une question d'une autre nature, conditionnelle :

- « si la CPFL se trouvait ailleurs, auriez-vous choisi une autre destination ou seriez-vous tout de même venus ici ? ».

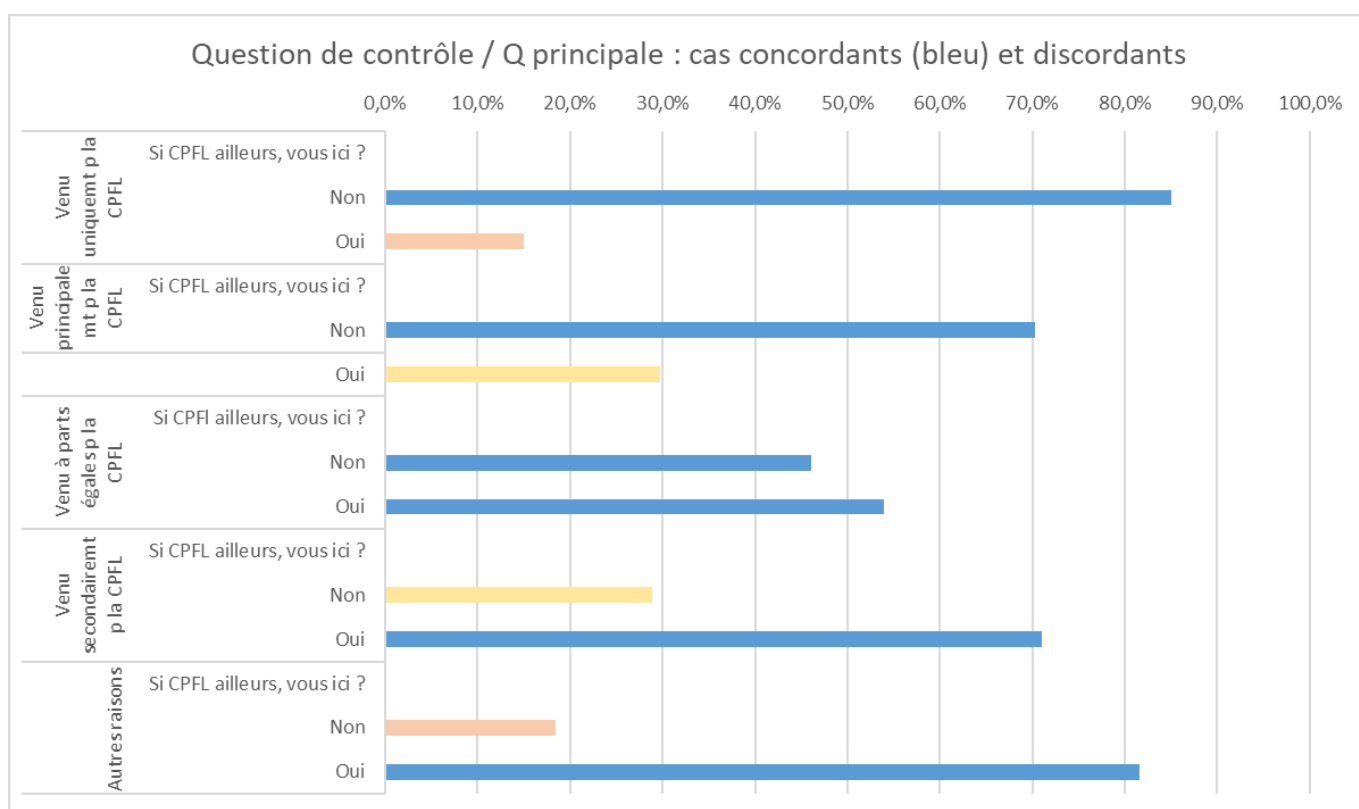
Cette forme de questionnement est celle de l'approche de substitution mais elle n'est utilisée que pour corriger les réponses à la question principale en cas de discordance.

Les cas de discordance sont à distinguer en dehors de la réponse « à parts égales » pour laquelle l'enquêté est en quelque sorte « forcé » à prendre parti par la question dite de contrôle. Le total des autres discordances est alors de 5+5+3+3 % de l'effectif total soit 16% ce qui est la fois significatif et nettement minoritaire.

De plus on observe qu'ils sont moins nombreux dans les cas « tranchés » (effet de 1 ou de 0) : si l'on se réfère à présent aux % au sein du cas considéré on obtient une répartition « en cloche » : 15%, 30% 29% 18%.

| Vous êtes venus ... | | | |
|--------------------------------------|---|--------------------|---------------|
| | Si CPFL était ailleurs seriez-vous tout de même ici ? | | |
| Venu uniquement p la CPFL | Si CPFL ailleurs, vous ici ? | % motif eff. redr2 | % total eff.r |
| | Non | 85% | 33% |
| | Oui | 15% | 5% |
| Venu principalemt p la CPFL | Si CPFL ailleurs, vous ici ? | % | |
| | Non | 70% | 14% |
| | Oui | 30% | 5% |
| Venu à parts égales p la CPFL | Si CPFL ailleurs, vous ici ? | % | |
| | Non | 46% | 6% |
| | Oui | 54% | 8% |
| Venu secondairemt p la CPFL | Si CPFL ailleurs, vous ici ? | % | |
| | Non | 29% | 3% |
| | Oui | 71% | 9% |
| Autres raisons | Si CPFL ailleurs, vous ici ? | % | |
| | Non | 18% | 3% |
| | Oui | 82% | 14% |

Les cas discordants apparaissent en couleurs orange dans le graphique suivant :



Complément méthodologique :

On pourrait également recouper d'autres informations venant de :

- Question(s) sur l'influence de la CPFL dans les choix, comme ici : « *l'inscription Unesco a-t-elle influencé votre choix de destination ?* ». Quelques réponses discordantes avec le coefficient C1 auraient pu permettre une correction, complémentairement à C2.
- Question(s) sur la fréquence, également utilisée par certains auteurs, ici : « *combien de fois viendrez-vous dans la CPFL au cours de votre séjour ?* », en gardant à l'esprit la fiabilité relative des réponses, notamment lorsque l'entretien est en début de séjour.
- Question(s) sur les activités menées, ici : « *qu'aurez-vous fait au cours de tout votre séjour ?* ». Cela nécessiterait de poser des questions suffisamment fermées sur les pratiques, notamment pour permettre leur localisation (éviter de simples réponses telles que : « se promener » « randonner » etc.)

5. Modulation de C1 : coefficient C2

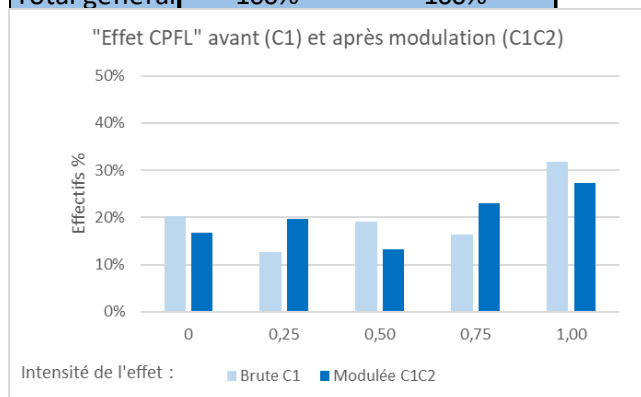
Méthode : dans ces cas discordants la position du coefficient dans l'échelle est *déplacée d'une unité en plus ou en moins* (+- 0,25), selon la réponse à la question de contrôle :

- si la réponse est « oui je serais (tout de même) venu ici » : discordance avec les C1 de 0,5 à 1, alors $C1C2 = C1 - 0,25$
- si la réponse est « non je ne serais pas venu ici » : discordance avec les C1 de 0,5 à 0, alors $C1C2 = C1 + 0,25$

Le résultat en est une diminution des extrêmes (1 et 0), d'autre part le report des certains cas discordants des tranches 0,75 et 0,25 sur la tranche 0,5 ne compense pas l'éclatement des réponses initialement à 0,5, éclatement vers 0,25 ou 0,75 du fait de la réponse à la question de contrôle. Cette tranche 0,5 a au final un effectif plus faible qu'initialement (13% du total) ⁵².

⁵² NB : dans cette étape de modulation, les 28 cas de non-réponse ont été traités en leur attribuant la position 0,5. Ce fait entraîne, dans le tableau qui suit, des proportions pour C1 légèrement différentes de celles portées dans les tableaux précédents.

| Touristes : Effet CPFL C1 et C1C2 | | |
|-----------------------------------|-------------|--------------|
| Intensité | Brute C1 | Modulée C1C2 |
| 1,00 | 32% | 27% |
| 0,75 | 16% | 23% |
| 0,50 | 19% | 13% |
| 0,25 | 13% | 20% |
| 0 | 20% | 17% |
| Total général | 100% | 100% |



Complément méthodologique, discussion :

D'autres méthodes de modulation/correction sont possibles :

- par exemple on pourrait considérer que la réponse à la question de contrôle a le même poids que la réponse à la question principale, bien qu'elle soit posée en second. Il reste à coder cette réponse par deux chiffres, par exemple :
C2contrôle oui = 0,25 et C2contrôle non = 0,75.

Puis on peut faire la moyenne entre C1 et C2contrôle :

$$C2var = (C1 + C2contrôle) / 2$$

Cette méthode a été testée : du fait de la relative symétrie des réponses dans ce cas particulier de la CPFL en 2019, le résultat final sur le calcul d'effet économique est **extrêmement proche** : pour une fréquentation de 800 000 visiteurs l'impact économique selon C2var serait de 69,4 Me, contre 68,9 selon la méthode retenue ici (voir infra en section III 3.1.3).

Il reste légitime de poursuivre une discussion sur le choix de la meilleure méthode pour plus de généralité. Concernant la question principale, le fait de poser celle-ci en deux temps a sans doute fluidifié les réponses (moins d'hésitation devant un tableau à 5 positions) mais elle peut comporter des biais si l'enquête surestime les deux premières modalités (« *totalemment* » ou « *pas du tout* »), sans faire attention à la troisième (« *en partie* ») avec dans ce cas un biais possible d'assentiment (le Oui favorisé par rapport au Non).

En tout état de cause, poser la question dite « principale » en deux temps ou en une seule fois ne peut que donner des résultats différents. On peut envisager, dans une perspective d'observation dans la durée, si une enquête ultérieure est menée en face à face et si le questionnement en deux temps est conservé (ne serait-ce que pour garder la continuité avec

2019), de consolider les réponses en récapitulant la réponse en deux temps par une vérification synthétique sur l'échelle à 5 valeurs. Cela permettrait de tester les écarts et de traiter les deux résultats, l'un pour une étude d'évolution, l'autre pour une estimation nouvelle.

Par ailleurs la question dite « de contrôle » donne-t-elle les mêmes résultats selon qu'elle est posée avant ou après la première ?

Tous ces points seront repris en troisième partie du rapport, partie consacrée à la perspective de méthode simplifiée et reproductible. On peut déjà souligner que l'objectif d'observer les évolutions du résultat économique au cours du temps donne *une priorité à la stabilité de la méthode* quelle que soit l'option retenue.

B. Les activités dénombrables des visiteurs : contribution à la mesure de la fréquentation

L'estimation d'un effet global de la fréquentation de la CPFL nécessite de disposer d'une estimation globale de cette fréquentation. Comme la plupart des variables utiles (notamment de dépenses de séjour, de transport, etc.) sont liées au visiteur et non à une visite ponctuelle, cette estimation doit être élaborée en nombre de visiteurs.

Cette question est traitée par une autre tâche du projet Valo-sites (T1-A), tâche qui évalue les divers dispositifs existants ou potentiels de dénombrement. En parallèle, sans préjuger des conclusions de T1-A il a été décidé de contribuer à cette estimation au sein de T1-B notamment en mobilisant deux dispositifs de dénombrement (« billetteries » et « éco-compteurs ») et certains résultats de l'enquête auprès des visiteurs.

1. Principe de l'estimation d'une fréquentation-plancher de la CPFL en visiteurs (et non en visites), à partir du nombre de visites

La combinaison des données issues des billetteries et des « éco-compteurs », moyennant consolidation de ces données autant que possible, fournit une estimation du nombre de visites sur les sites dits « à dénombrement ». C'est ce qui a permis d'établir a priori avant l'enquête un plan d'échantillonnage.

Il s'agit ici d'estimer un nombre de *visiteurs* de la CPFL simple addition de données de *visites* n'élimine pas les « doubles comptes », c'est à dire les décomptes excessifs provenant du fait qu'un même visiteur peut être vu dans plusieurs de ces sites. Ces derniers doivent être corrigés pour bien estimer la fréquentation plancher de la CPFL.

L'enquête permet d'éliminer les doubles-comptes entre éco-compteurs, entre billetteries et entre billetteries et éco-compteurs, puisqu'elle permet de connaître les proportions de visites *multiples* (nombre de visiteurs multiplié par nombre de sites) sur la CPFL. C'est une méthode originale qui est proposée et appliquée ici.

Dans ce qui suit on parle de « fréquentation-plancher » du simple fait que les billetteries et les éco-compteurs ne comptabilisent pas toutes les fréquentations de la Chaîne des Puys. Ce point est traité infra en section IV.2.3

2. Calcul des taux de comptes multiples

Les taux sont calculés à partir de l'enquête auprès des visiteurs. Les sites visités retenus dans le calcul des taux de comptes multiples sont ceux de la CPFL dont les données de fréquentation sont disponibles soit via les éco-compteurs, soit via les billetteries ou autres sources de comptage. Les données de billetteries des sites payants (Panoramique, Vulcania, Lemptégy, Grotte de la pierre), les données de la fréquentation de la maison du PNR et de Volvic et celles des éco-compteurs situés sur les sentiers menant à certains puys, peuvent permettre de calculer cette fréquentation plancher⁵³.

Suppression des « doubles comptes »

À partir des données de ce type d'enquête on peut faire une estimation des « doubles comptes ». Ceux-ci sont dus au fait que certains visiteurs passent dans une, dans deux ou dans les trois entreprises payantes de la CPFL, que certains (les mêmes ou d'autres) visitent un ou plusieurs puys dotés d'éco-compteurs sur le site d'étude donc sont potentiellement passés devant un de ces éco-compteurs. Enfin certains sont donc à la fois passés devant des éco-compteurs et ont visité un ou des sites payants. L'enquête permet de faire une estimation de ces « doubles comptes » par le fait que les visiteurs ont déclaré, dans les enquêtes les plus complètes (la moitié de l'échantillon environ), toutes les activités pratiquées au cours de leur journée et de leur séjour (cas des touristes), donc notamment les activités citées ci-dessus⁵⁴.

Estimation de fréquentation plancher

La méthode est la suivante : supposons que dans notre échantillon d'enquête p_s (en %) de visites sont effectuées sur s sites (s entier naturel non nul) prenant toutes les valeurs de 1 à environ 15 selon le nombre de sites (avec éco-compteurs ou billetteries ou données de fréquentations disponibles exploitables). On dispose de toutes les combinaisons fiables pour au moins les « enquêtes longues ». Pour éliminer les doubles comptes, on procède ainsi :

- appliquer au total des visites comptabilisées (par les billetteries, les éco-compteurs, les données de fréquentations disponibles) les pourcentages p_s (en %) constatés dans l'échantillon enquêté de visiteurs. Cela donne des sous-populations de visites, une par valeur de s .

⁵³ Les éco-compteurs du Puy de Côme et du Puy Nugère ne sont pas pris en compte car ces puys ne sont pas visités dans notre base. Nous n'avons pas utilisé les données des mois de Mars, Avril et Novembre puisque nous n'avons pas pu effectuer des enquêtes pendant ces périodes. Aussi, les taux calculés sont issus des enquêtes de trois (3) enquêteurs sur cinq (5). En effet, pour des raisons diverses (début de l'enquête, effet-enquêteur, etc.), deux enquêteurs n'ont pas nécessairement noté tous les sites visités.

⁵⁴ On verra plus loin que l'enquête doit couvrir la totalité des activités pratiquées par les visiteurs interrogés au cours de l'année, ce qui est partiellement le cas ici, du fait que les excursionnistes répétés (ou les touristes répétés) n'ont pas été interrogés sur le détail de leurs activités pendant les autres venues ou séjours. Des estimations ont dû être faites (cf infra).

- Puis diviser les sous-populations de visites par le nombre s de sites visités qui concernent chacune, donne des sous-populations de visiteurs, une par valeur de s .
- Enfin additionner ces dernières permet d'obtenir une population totale estimée de visiteurs de la CPFL (valeur-plancher faute d'autres mesures de fréquentation)

Ce qui revient à faire le calcul suivant. Soit :

V = Nombre total de visites sur la CPFL y compris les doubles-comptes

N_s = Nombre total de visiteurs de la CPFL visitant s sites

N = Nombre total de visiteurs de la CPFL

p_s (en %) = Pourcentage de visites effectuées sur s sites dans notre échantillon d'enquête

On obtient :

$$N_s = \frac{V * p_s}{s} \quad \text{et} \quad N = \sum_{(s=1...n)} N_s$$

3. Le cas des visiteurs non dénombrables

Il existe une grande diversité de pratiques ouvertes aux visiteurs du site Chaîne des Puys–Faille de Limagne. Un grand nombre de celles-ci n'entraînent pas de décompte, ni par billetterie, ni pas passage devant un éco-compteur. On rappelle que les éco-compteurs ont pour objet prioritaire de connaître les fréquentations qui entraînent des risques pour le milieu, seuls les sites les plus fréquentés et fragiles sont équipés.

L'enquête n'était pas conçue pour pallier le manque de connaissance quantitative sur la fréquentation. Seuls les principaux parkings officiels ont été retenus ainsi que les sites de visite (avec billetterie ou dénombrement).

Pour des raisons logistiques les lieux d'enquête ne pouvaient pas être dispersés en tous points où des véhicules peuvent se garer. De plus concernant le Puy de Dôme les enquêtes qui se sont déroulées à l'intérieur de la gare de départ du train présentaient beaucoup d'avantages sur les enquêtes menées sur parking (abri pendant la canicule 2019, périodes d'attente des départs de train favorables à une enquête, etc.), ce qui a conduit à ne pas enquêter sur le parking correspondant, or ce parking important peut desservir aussi des randonnées hors éco-compteur (sentier du tour du Puy de Dôme).

Ainsi une limite de l'enquête⁵⁵ est d'avoir sous-échantillonné :

⁵⁵

- Les parkings secondaires (par exemple au départ d'un itinéraire dans la Cheire de Come, etc.)
- les points de stationnement dit « sauvage », dont certains peuvent accueillir une dizaine de véhicules et plus,
- de même que certaines pratiques sportives (cyclistes avec point de départ hors-zone, etc.)

Sur les lieux d'enquête certains visiteurs rencontrés ont déclaré n'avoir pas fréquenté de site dénombrable : sur tout parking il est possible de pratiquer un sentier qui ne monte pas au sommet du ou des Puys correspondants.

3% des visiteurs enquêtés sont dans ce cas, sachant que les touristes ont déclaré leurs activités sur tout le séjour ce qui réduit cette proportion par rapport à la situation du jour d'enquête. Ces 3% environ ont été ajoutés aux effectifs globaux après le calcul exposé ci-dessous. Mais cet ajout ne comble pas le déficit exposé dans la section précédente lié au fait que tous les lieux de présence et d'activité n'ont pu être enquêtés.

4. Le cas des « excursionnistes répétés »

Il était demandé aux excursionnistes à quelle fréquence ils venaient sur le site au cours de l'année. Leurs réponses ont été prises en compte, en pratiquant deux modulations :

- Sur le nombre d'excursions : les réponses étaient des estimations globales sans réflexion sur les cas particuliers : ainsi un enquêté qui dira « quasiment tous les jours » n'a pas eu le temps de réfléchir à tous les aléas, climatiques ou de santé ou autres qui ont fait exception. Sur ce point un scénario est construit avec coefficients de réduction, scénario appelé Sc fr4, dont les coefficients sont : 0,9 sur "tous les mois", 0,75 sur "toutes les semaines", 0,5 pour les deux cas avec réponses "trois quatre fois par semaine", et "presque tous les jours".
- Sur le nombre moyen de « sites à dénombrement » visités par excursion : l'enquête n'est pas allée jusqu'à leur demander combien de fois ils étaient passés sur des points « à dénombrement » (sentiers menant aux sommets, ou sites avec billetterie). Il a fallu faire une estimation avec un coefficient dit Sc den x (pour « Scénario dénombrement x »). Le coefficient retenu ici correspond à l'incertitude, soit une probabilité 0,5 d'être dénombré à chaque excursion ultérieure.

NB : ces excursionnistes répétés apparaissent en haut de l'échelle en nombre de visites dénombrables, échelle qui s'étend très largement au-delà des chiffres 0 à 7 trouvés dans les autres réponses (le tableau qui suit présente l'échelle de 0 à 37), mais il faut se souvenir que le redressement probabiliste rétablit le fait que les effectifs de ces excursionnistes dans la population totale sont réduits.

5. Résultat : estimation d'une fréquentation-plancher

Un calcul simple part de l'hypothèse que l'échantillon redressé est représentatif de la population totale dénombrable. Cet échantillon « porte » un nombre v de visites en sites dénombrables pour un effectif redressé n . Le nombre de visites dénombrées dans la population totale étant noté V on obtient par une « règle de trois » :

$$N_p = n * V / v$$

Les valeurs $n = 1\ 734$ $v = 2\ 699$ et $V = 1\ 317\ 933$ donnent un effectif plancher

$$N_p = 846\ 366 \text{ visiteurs}$$

Une autre façon plus détaillée de mener ce calcul est indiquée dans le tableau suivant.

La colonne centrale « % visites dénombr. / total » exprime, en pourcentage du total, le produit des effectifs par le nombre de visites pour chaque classe de nombres de visites. *L'hypothèse de représentativité est ici que cette colonne calculée sur l'échantillon enquêté représente également les pourcentages au sein de la population totale des visiteurs.*

En appliquant alors ces proportions au nombre total de visites dénombrées (1 317 933) on obtient le détail de la colonne « Nb visites dénombr. ds population ». E

Enfin en divisant chacun de ces nombres par le nombre de visites qui définit chaque classe, on obtient l'effectif de visiteurs pour chaque classe et la somme donne la fréquentation-plancher totale (après avoir ajouté les 3% environ, déduits des « visiteurs non dénombrables » rencontrés lors de l'enquête).

| Estimation-plancher de l'effectif global des visiteurs de la CPFL à partir du nombre de visites | | | | | | | | |
|---|---------------------|---------------------|-----------------------------------|----------------------------|-----------------------------------|------------------------------|-------|---------------------------------------|
| Nbre sites dénombr. visités | Effectif de groupes | % effectifs / total | Nbres visites dénombr. Ds échant. | % visites dénombr. / total | Nb visites dénombr. ds population | Eff. visiteurs ds population | % | Eff. visiteurs + sans visite dénombr. |
| Excurs. | 486,63 | 28,08% | 787,1 | 29% | 383 376 | 221 110 | 27,3% | 229 262 |
| 0 | 16,90 | 0,97% | - | 0% | 0 | ? | 0 | |
| 1 | 345,13 | 19,92% | 343 | 12,7% | 167 314 | 167 314 | 20,7% | |
| 2 | 55,87 | 3,22% | 93 | 3,5% | 45 491 | 22 746 | 2,8% | |
| 3 | 31,51 | 1,82% | 84 | 3,1% | 40 675 | 13 558 | 1,7% | |
| 4 | 10,15 | 0,59% | 36 | 1,3% | 17 478 | 4 369 | 0,5% | |
| 5 | 5,53 | 0,32% | 27 | 1,0% | 13 380 | 2 676 | 0,3% | |
| 6 | 14,74 | 0,85% | 88 | 3,3% | 42 843 | 7 141 | 0,9% | |
| 7 | 1,37 | 0,08% | 10 | 0,4% | 4 659 | 666 | 0,1% | |
| 8 | 0,22 | 0,01% | 2 | 0,1% | 812 | 101 | 0,0% | |
| 19 | 1,49 | 0,09% | 28 | 1,0% | 13 773 | 725 | 0,1% | |
| 20 | 3,36 | 0,19% | 67 | 2,5% | 32 754 | 1 638 | 0,2% | |
| 21 | 0,25 | 0,01% | 5 | 0,2% | 2 588 | 123 | 0,0% | |
| 23 | 0,05 | 0,00% | 1 | 0,0% | 574 | 25 | 0,0% | |
| 37 | 0,06 | 0,00% | 2 | 0,1% | 1 036 | 28 | 0,0% | |
| Touristes | 1 246,29 | 71,92% | 1 919 | 70,9% | 934 557 | 588 145 | 72,7% | 606 890 |
| 0 | 38,85 | 2,24% | - | 0,0% | - | - | 0,0% | |
| 1 | 676,10 | 39,02% | 676,10 | 25,0% | 329 329 | 329 329 | 40,7% | |
| 2 | 400,52 | 23,11% | 801,03 | 29,6% | 390 182 | 195 091 | 24,1% | |
| 3 | 92,45 | 5,34% | 277,36 | 10,3% | 135 101 | 45 034 | 5,6% | |
| 4 | 30,13 | 1,74% | 120,50 | 4,5% | 58 696 | 14 674 | 1,8% | |
| 5 | 6,58 | 0,38% | 32,89 | 1,2% | 16 022 | 3 204 | 0,4% | |
| 6 | 0,96 | 0,06% | 5,73 | 0,2% | 2 793 | 465 | 0,1% | |
| 7 | 0,71 | 0,04% | 5,00 | 0,2% | 2 433 | 348 | 0,0% | |
| Total généra | 1 732,92 | 100,00% | 2 705,67 | 100,0% | 1 317 933 | 809 255 | 100% | 836 152 |

L'estimation-plancher est ici de $N_p = 836\ 152$ visiteurs, mais avec une marge d'erreur dûe aux arrondis opérés pour chaque tranche en nombre de visites.

Compte-tenu de l'incertitude sur plusieurs points qui ont nécessité des ajustements hypothétiques (cf ci-dessus) et de l'incertitude sur les visiteurs « non-dénombrables non enquêtés », qui s'ajoutent à cette estimation, et en attente d'estimation éventuelle par d'autres dispositifs :

une fourchette de 800 000 à 900 000 visiteurs est proposée, et les résultats économiques seront présentés selon ces deux valeurs.

a) Le cas des « touristes » répétés

De même que les excursionnistes, les touristes peuvent venir plusieurs fois en séjour dans l'année et visiter la CPFL au cours de ces séjours. On les appellera les « touristes répétés ».

Ces touristes répétés sont très minoritaires : le SDT fourni par Kantar ex-Sofres pour le Département 63 en 2018 les estima à 11% de l'ensemble hors voyages professionnels. De plus

l'hypothèse a été faite que ces séjours répétés ont d'autres motifs que la seule attractivité de la CPFL (plus souvent ce serait la visite d'amis ou de famille). Enfin le questionnaire ne pouvait comporter toutes les questions nécessaires à une évaluation précise, notamment sur les activités et dépenses des autres séjours.

Un simple test est effectué de l'effet d'une imputation de quelques sorties supplémentaires en CPFL pour 9% environ des touristes interrogés. Un tirage aléatoire de ces 9% de touristes est effectué et 1, 2 ou 3 sorties en sites dénombrables sont introduites pour ces touristes répétés, pour tester l'effet sur l'estimation-plancher du nombre de visiteurs et de la répartition excursionnistes/touristes.

L'effet attendu est une baisse des effectifs globaux des visiteurs du fait que l'échantillon pour un même effectif « porte » à présent un nombre plus élevé de sorties dénombrées. Cet effet est réel mais faible dans la mesure où l'on reste autour de 800 000 visiteurs alors que c'est au final une fourchette de 800 à 900 000 visiteurs qui est retenue comme exposé ci-dessus du fait de l'incertitude bien supérieure sur les visites non dénombrées (visites diffuses).

| Test de prise en compte des séjours répétés ("touristes répétés") | | | | |
|---|--------------------------------------|--|--------------------------------|-------------|
| | nb de groupes de Touristes répétés * | nombre de visites dénombrables suppl. (hypothèses) | Effectifs globaux de visiteurs | % touristes |
| référence | 0 | 0 | 838 358 | 72,6% |
| test 1 | 64 | 1 | 814 506 | 71,8% |
| test 2 | 64 | 2 | 804 791 | 71,5% |
| test 3 | 64 | 3 | 790 931 | 71,3% |

Concernant l'effet économique supplémentaire de ces séjours répétés il n'est pas chiffrable faute de détails mais, l'impact économique strict *unitaire* (un touriste répété porte davantage de dépenses à l'année que ce qu'il a déclaré pour le séjour lors de l'enquête) est faible si nous maintenons l'hypothèse que la motivation CPFL du touriste répété (coefficient C2) est faible.

Et au niveau global la baisse des effectifs joue dans le sens contraire de l'effet économique supplémentaire, ce qui est attendu. Ainsi au final on se trouve devant une des composantes de l'intervalle de confiance global difficile à calculer, mais on peut estimer que l'impact économique unitaire supplémentaire (faible) ne fait que compenser la (faible) baisse d'effectifs dénombrables.

Conformément à la section « III.3.7 Imputer ou simuler ? » la BDD n'a pas été modifiée car des imputations seraient aléatoires, mais les différences en effectifs et taux de

touriste/excursionnistes issues de la simulation sont prises en compte dans l'appréciation des intervalles de confiance.

C. Les dépenses des visiteurs : premier aperçu

Les dépenses des visiteurs après traitements de base (vérifications, corrections éventuelles, imputation de valeurs absentes) sont donc disponibles avec un détail par groupe enquêté et par poste de dépense.

Avant d'exposer en section V l'estimation des effets économiques liés à ces dépenses, deux éclairages sont donnés ici, sur : 1- La hiérarchie des dépenses, 2- La comparaison avec une autre source d'information.

1. Hiérarchie des dépenses

Au sein de l'échantillon, les valeurs en pourcentage éclairent la hiérarchie des dépenses par poste et par type de visiteur, celle-ci s'établit ainsi :

| Ventilation des dépenses au sein de l'échantillon redressé | | | | | | |
|--|----------------|-------------|-------------------|-------------|----------------|-------------|
| Montants et % par poste | touristes | | excursionnistes * | | Total | |
| Postes de dépense | Montant | % | Montant | % | Montant | % |
| Hébergement | 107 240 | 38% | - | 0% | 107 240 | 35% |
| Restaurant | 59 213 | 21% | 6 297 | 22% | 65 510 | 21% |
| Alimentation | 31 470 | 11% | 1 567 | 5% | 33 037 | 11% |
| Achat de souvenirs | 20 807 | 7% | 1 846 | 6% | 22 653 | 7% |
| Activités payantes | 35 329 | 13% | 13 466 | 47% | 48 796 | 16% |
| Services | 7 476 | 3% | 124 | 0% | 7 600 | 2% |
| Transports locaux | 20 090 | 7% | 5 487 | 19% | 25 578 | 8% |
| Total | 281 626 | 100% | 28 787 | 100% | 310 413 | 100% |

Le poste *d'hébergement* représente de loin le principal poste de dépense, même pour l'ensemble des visiteurs puisqu'il pèse 35% du total, devant le poste de la *restauration* (21%) suivi par celui des *activités payantes* (16%). Viennent enfin les dépenses de *transport local* et les *achats de souvenirs et produits locaux* qui ne sont pas négligeables avec respectivement 8% et 7% du total.

La structure des dépenses est bien différente entre touristes et excursionnistes, d'abord du fait de l'absence d'hébergement pour les seconds.

Une analyse plus détaillée sera présentée en Partie II du rapport. Elle concernera surtout les touristes qui occasionnent l'essentiel des flux monétaires, et qui sont, avec les quelques excursionnistes non-locaux, les seuls fournisseurs de l'impact économique local au sens strict (cf infra section V). De plus il faut rappeler que les dépenses des excursionnistes souffrent d'une approximation particulière due aux incertitudes sur les excursions répétées (localisations et activités).

2. Comparaison avec les données du SDT (Suivi de la Demande Touristique, INSEE-Kantar)

Le SDT est une production annuelle de données sur les caractéristiques et les dépenses de voyage et de loisirs des français. Cette production se fonde sur les réponses d'un panel de 22 000 français interrogés chaque mois. Le panel est géré par le bureau d'études Kantar (ex-TNS-Sofres).

La production principale est de niveau national, mais Kantar fournit des résultats à une échelle plus fine, généralement régionale, parfois départementale à la demande. Le Département du Puy de Dôme a acquis les données et l'analyse à son échelle pour l'année 2018. Du fait de la taille du panel (818 voyages décrits en 2018 pour le Puy de Dôme) l'outil ne convient pas pour une échelle plus proche de celle de la CPFL. Par ailleurs les 818 voyages recouvrent tout l'éventail des motivations (professionnelle, personnelle) ; parmi les motivations personnelles, majoritaires, Kantar distingue, à l'échelle de Département, et exprimé en voyages (non en durées) : « vacances loisirs : 47,9%, visite famille amis : 42,3%, santé : 1,5%, autres : 8,1%. En partie II infra, la présentation des caractéristiques des visiteurs de la CPFL montrera des différences en matière de profils, de caractéristiques des séjours, etc.

La comparaison entre les résultats de l'enquête Valo-sites pour la seule CPFL en 2019 et les données SDT pour tout le Puy de Dôme en 2018 a *toutes raisons de montrer des différences*, mais a semblé tout de même utile pour vérifier l'absence ou non d'anomalies. En partie II infra, on verra des différences en matière de profils, de caractéristiques des séjours, etc. Ici on comparera les montants de dépense des touristes, par personne et par nuit, et selon les modes d'hébergement choisis. Ces filtres réduisent d'autant les sources de différences à expliquer. Précision méthodologique :

- Comme ceux de Kantar les chiffres ci-dessous sont calculés en montant de dépense totale par adulte et par nuit.
- Les chiffres du SDT indiqués sont ceux de la catégories « voyages personnels » c'est-à-dire hors voyages professionnels
- le panel SDT est français, nous avons exclu les réponses de personnes étrangères ou habitant à l'étranger
- Il semble que dans le SDT, les montants moyens aient été calculés en « moyennes des moyennes » (moyenne sur les 818 voyages des moyennes de dépense par adulte pour chaque voyage), les résultats Valo-sites sont exceptionnellement calculés ici sur le même mode, en notant que la méthode « somme sur somme » adoptée par Valo-sites (ici : somme des dépenses sur somme des adultes) donne des résultats systématiquement inférieurs, ainsi que c'est confirmé dans la littérature.

La comparaison avec le SDT donne ainsi :

| Touristes français : dépenses totales par adulte et par nuit selon les types d'hébergement | | | | |
|---|------------|-----------------------------|--------------------------------------|----------------------|
| SDT 63 2018 | | Valo-Sites CPFL 2019 | | adultes*nuits |
| hôtel | 144 | 134 | hôtel | 591 |
| « Location, Gîte rural et Chambre d'hôte » | 74 | 66 | Gîte, location à un particulier | 1 254 |
| | | 85 | « Airbnb » | 247 |
| | | 96 | Chambres d'hôtes | 86 |
| | | 77 | Gîte, Airbnb, Chambre d'hôtes | 1 587 |
| (où sont les Mobile Homes ?) | 57 ou 74 ? | 76 | Camping-mobile-home | 350 |
| Camping | 57 | 48 | Camping-car | 440 |
| | | 51 | Camping : tente, caravane | 292 |
| Famille | 43 | 39 | Chez de la famille, des amis | 665 |
| Amis | 67 | | | |
| « ST non marchand » | 39 | 28 | Résidence secondaire | 275 |
| « autre hbgt payant » | 102 | 77 | Village / club vacances | 110 |

Les valeurs sont différentes, ce qui était attendu, mais assez proches, ce qui conforte le présent travail.

Il restera à interpréter le fait que, en 2019, calculées avec la méthodologie Valo-sites, les dépenses des visiteurs de la CPFL semblent (légèrement) *inférieures* (*partout sauf* dans les cas « gîte, location, airbnb, chambre d'hôtes »), à celles des visiteurs de l'ensemble du Département 63, en 2018, avec la méthodologie SDT.

V. CPFL : mesure de l'activité économique générée par la visite du site

Les parcs nationaux, régionaux et les sites naturels en général apportent une contribution significative à l'économie des zones (régions, départements, communes) dans lesquelles ils sont situés. La grande majorité des études mesurant l'activité économique générée par la fréquentation d'un site se concentrent sur les dépenses des visiteurs en utilisant des méthodes d'enquête. Elle repose sur l'idée que les dépenses effectuées par les visiteurs suscitent un certain volume d'activité économique qui crée des revenus et procure des emplois aux populations locales. L'approche habituelle consiste à estimer des moyennes de dépenses qui peuvent être appliquées à la fréquentation du site pour calculer les dépenses totales, qui peuvent à leur tour être appliquées à un modèle d'impact économique ou à un ensemble de multiplicateurs.

Dans cette partie du rapport, l'activité économique générée par la visite de la CPFL est mesurée à partir des dépenses touristiques directes dans les départements du Puy-de-Dôme où elle se trouve et dans les départements de l'Allier, du Cantal et de la Haute-Loire (les caractéristiques des quatre départements à trouver en Annexe 1). Cette activité est appréhendée au travers de deux types d'analyse : l'importance économique et l'impact économique dans chacune des quatre départements (Puy-de-Dôme ; Allier ; Cantal et la Haute-Loire). L'importance économique est estimée à partir de la dépense totale de tous les visiteurs. L'impact économique de son côté peut être estimé à partir de trois catégories : impacts directs, impacts indirects et impacts induits (paragraphe II.2.2). Cependant, l'indisponibilité des outils statistiques au à l'échelle départementale permettant d'estimer les impacts indirects conduit à ne considérer que les impacts directs et induits.

Les deux types d'analyse se reposent sur la combinaison d'informations en provenance de trois types de sources statistiques (figure 2) :

- Les bases de données des Eco-compteurs et des billetteries, qui fournissent des informations sur la fréquentation des sites
- Les enquêtes « visiteur » réalisées dans le cadre du projet, qui fournissent des informations sur les comportements de dépense des visiteurs de la CPFL ;
- Les bases de données de l'INSEE, qui fournissent des informations sur la démographie, l'emploi et les revenus à l'échelle locale.

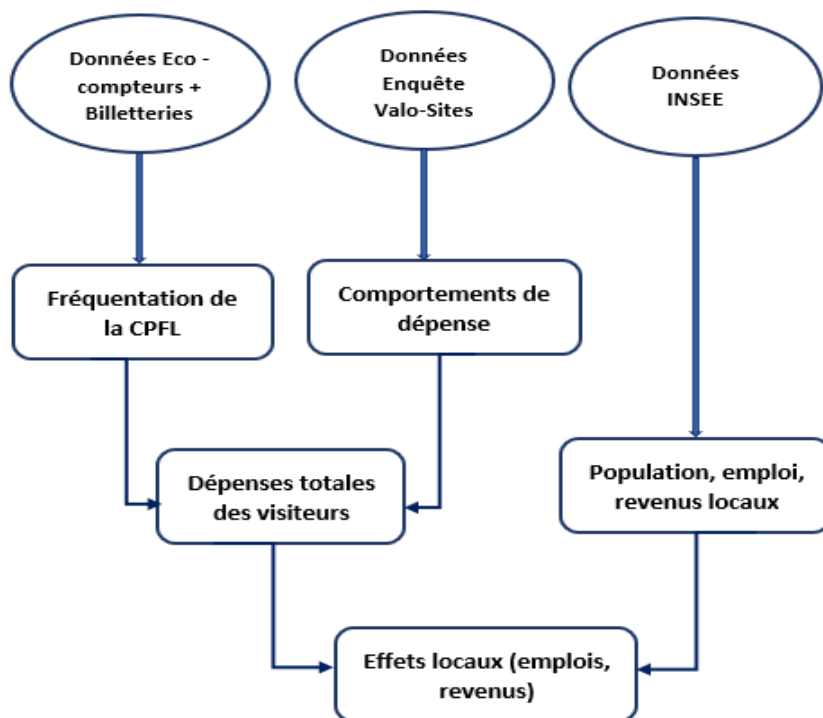


Figure 3 : Sources des informations utilisées pour estimer les effets locaux de la visite de la CPFL

A. Détermination des impacts directs

1. Localisation des dépenses

L'impact direct de la fréquentation de la CPFL sur les zones d'étude (les départements du Puy-de-Dôme, de l'Allier, du Cantal et de la Haute-Loire) correspond à la dépense totale des visiteurs réalisée dans chaque zone. Cette dépense totale est déterminée à partir d'une enquête socio-économique auprès des visiteurs de la CPFL (**section III.2**). L'analyse de l'activité économique se faisant par rapport à une zone déterminée, il est primordial d'identifier le lieu de réalisation de chaque poste de dépense. Malheureusement, des questions ne pouvaient pas être posées sur les lieux de réalisation exacte de chaque poste de dépense. En effet, ces questions alourdisent le questionnaire et demandent un effort supplémentaire aux personnes enquêtées.

Ainsi donc, une hypothèse sur le lieu de réalisation de chaque type de dépense a été faite. A cet effet, en ce qui concerne les touristes, les dépenses d'hébergement, de restauration, d'alimentation, de service à la personne et la moitié des dépenses de transport et de produits locaux et souvenirs sont affectées à leur lieu d'hébergement. Le reste des dépenses (activités transport, produits locaux et souvenirs) étant attribué au département des activités (ici le Puy-de-Dôme). Du côté des excursionnistes, les dépenses liées au service à la personne et la moitié des dépenses de transport sont attribuées au département d'habitation. Le reste des dépenses étant attribué au département des activités. **Le tableau 8** récapitule la localisation des postes de dépenses par type de visiteurs.

Tableau 5 : Localisation des dépenses

| Lieu des dépenses | Dépenses des Touristes | Dépenses des excursionnistes |
|---------------------------------|------------------------------------|------------------------------|
| Département d'hébergement | Hébergement | |
| | Restaurant (50%) | |
| | Alimentation | |
| | Service à la personne | |
| | Transport (50%) | |
| | Produits locaux et souvenirs (50%) | |
| Département d'habitation | | Transport (50%) |
| | | Service à la personne |
| Département des activités (PDD) | Activités | Restaurant |
| | Restaurant (50%) | Alimentation |
| | Transport (50%) | Activités |
| | Produits locaux et souvenirs (50%) | Transport (50%) |
| | | Produits locaux et souvenirs |

2. Affectation de tout ou partie des dépenses des visiteurs à « l'effet CPFL » (rappel)

Si les études d'impact économique du tourisme sont courantes, la validité des résultats dépend de diverses questions méthodologiques. Nombre de ces questions concernent les détails des méthodes de calcul des dépenses totales directes. En effet, sans mesures appropriées des dépenses directes des touristes liées aux activités touristiques, même un modèle d'impact économique ou un multiplicateur le plus détaillé et théoriquement approprié donnera des résultats trompeurs (Tyrrell et Johnston, 2001).

Afin de déterminer la proportion de ces dépenses totales qui pourrait être directement attribuable à la CPFL, un coefficient d'affectation a été généré à l'aide des enquêtes auprès des visiteurs. En effet, les estimations de l'activité économique de la CPFL peuvent être surévaluées si on utilise les dépenses totales des visiteurs plutôt que la proportion des dépenses totales consacrée uniquement à la CPFL (facteur d'attribution) et si la fréquentation d'autres sites alternatifs par les visiteurs ne sont pas pris en compte (comportement de substitution).

La section III.2. a présenté la méthode détaillée et les résultats de ce calcul de coefficients C1 et C2.

3. Calcul des dépenses moyennes par visiteur

La dépense moyenne par personne pour chaque type de visiteur a été calculée comme suit :

$$\bar{D} = \frac{\sum_{i=1}^n D_i}{\sum_{i=1}^n t_i}$$

Où \bar{D} est la moyenne des dépenses par personne par catégories de visiteurs ; D_i représente la dépense totale de chaque groupe de visiteurs i , t_i est la taille de chaque groupe de visiteurs (nombre de personnes), n est le nombre total de groupes ou de réponses dans l'échantillon.

Selon la catégorie des visiteurs et le lieu de réalisation des dépenses, sont identifiées les catégories de dépenses suivantes :

Pour les excursionnistes :

Exc. Hab. le PDD = excursionniste habitant le Puy-de-Dôme.

Exc. Hab. hors PDD = excursionniste habitant hors du Puy-de-Dôme.

Exc. Hab. l'Allier = excursionniste habitant l'Allier.

Exc. Hab. le Cantal = excursionniste habitant le Cantal.

Exc. Hab. la Haute-Loire = excursionniste habitant la Haute-Loire.

Pour les touristes :

Tour. Héb. au PDD = touriste hébergé au Puy-de-Dôme.

Tour. Héb. hors PDD = touriste hébergé hors du Puy-de-Dôme.

Tour. Héb. dans l'Allier = touriste hébergé dans l'Allier.

Tour. Héb. dans le Cantal = touriste hébergé dans le Cantal

Tour. Héb. dans la Haute-Loire = touriste hébergé dans la Haute-Loire

Les tableaux ci-dessous montrent les dépenses par personne pour chaque département et pour chaque type de visiteurs.

Tableau 6 : *Dépenses moyennes par visiteur selon le type de visiteurs et le type de dépenses considérées (euros)*

| Catégorie de visiteurs | Dépenses totales | | | Dépenses dans le 63 attribuables à la CPFL | | | |
|------------------------|-----------------------------|---------------------------------|-----------------------------|--|-----------------------------|--------|-----------------------------|
| | Séjour ou jours d'excursion | Part dans la dépense dans le 63 | Part dans la dépense totale | C1 | Part dans la dépense totale | C2 | Part dans la dépense totale |
| Exc. Hab. hors PDD | 56,25 | 45,43 | 81% | 45,43 | 81% | 45,43 | 81% |
| Exc. Hab. le PDD | 60,92 | 60,92 | 100% | 60,92 | 100% | 60,92 | 100% |
| Tour. Héb. hors PDD | 190,54 | 40,42 | 21% | 21,03 | 11% | 20,23 | 11% |
| Tour. Héb. au PDD | 229,61 | 229,61 | 100% | 124,99 | 55% | 122,79 | 55% |

Tableau 7 : *Dépenses moyennes par type de visiteurs et le type de dépenses considérées dans les autres départements (euros)*

| Catégorie de visiteurs | Dépense par personne dans le département concerné | Dépense par personne affectable | | Dépense par personne affectable | |
|--------------------------------|---|---------------------------------|---|---------------------------------|---|
| | | (C1) à la CPFL | Part dans la dépense par personne dans le département | (C2) à la CPFL | Part dans la dépense par personne dans le département |
| Exc. Hab. l'Allier | 4,40 | 4,40 | 100% | 4,40 | 100% |
| Tour. Héb. dans l'Allier | 196,45 | 87,46 | 45% | 67,34 | 34% |
| Exc. Hab. le Cantal | 7,32 | 7,32 | 100% | 7,32 | 100% |
| Tour. Héb. dans le Cantal | 280,16 | 139,19 | 50% | 128,39 | 46% |
| Exc. Hab. la Haute-Loire | 9,84 | 9,84 | 100% | 9,84 | 100% |
| Tour. Héb. dans la Haute-Loire | 205,11 | 95,97 | 47% | 112,76 | 55% |

4. Calcul des dépenses des visiteurs dans chaque département : totaux, importances économiques, impacts économiques

a) Calcul principal

Les données de l'enquête ont permis d'estimer les dépenses moyennes par catégorie de visiteurs (excursionnistes et touristes). Les dépenses moyennes par personne ont ensuite été multipliées par le nombre total de chaque catégorie de visiteurs (tableau 11) pour calculer les dépenses totales dans chaque département.

Tableau 8 : **Fréquentation de la CPFL** : détail pour niveaux-planchers de 800 ou 900 000 visiteurs

| Répartition des excursionnistes dans la population (Détaillée) | Effectif redressé échantillon | % | Effectif population hyp. 800 000 | Effectif population hyp. 900 000 |
|--|-------------------------------|------|----------------------------------|----------------------------------|
| Exc hors PDD | 188 | 39% | 86 759 | 97 603 |
| <i>dont Allier</i> | 80 | 43% | 37 014 | 41 641 |
| <i>dont Cantal</i> | 19 | 10% | 8 977 | 10 099 |
| <i>dont Haute-Loire</i> | 10 | 5% | 4 631 | 5 210 |
| <i>dont Autre</i> | 78 | 42% | 36 136 | 40 653 |
| Exc dans PDD | 299 | 61% | 137 917 | 155 156 |
| Exc | 487 | 100% | 224 675 | 252 760 |
| Touriste hors PDD | 120 | 10% | 55 176 | 62 073 |
| <i>dont Allier</i> | 41 | 34% | 18 737 | 21 079 |
| <i>dont Cantal</i> | 15 | 13% | 7 141 | 8 034 |
| <i>dont Haute-Loire</i> | 9 | 7% | 3 965 | 4 460 |
| <i>dont Autre</i> | 55 | 46% | 25 332 | 28 499 |
| Touriste dans PDD | 1127 | 90% | 520 149 | 585 167 |
| Touriste | 1247 | 100% | 575 325 | 647 240 |
| Total | | | 800 000 | 900 000 |

Tableau 9 : Dépenses totales selon le type de visiteurs et le type de dépenses considérées (euros) : dans le PDD

(Hyp. 800 000 visiteurs)

| Hyp. 800 000 visiteurs | Dépenses totales | Dans le PDD | Dépenses dans le PDD attribuables à la CPFL | |
|---|--------------------|--------------------|---|-------------------|
| | | | C1 | C2 |
| Exc. Hab. hors PDD (1) | 4 879 838 | 3 941 063 | 3 941 063 | 3 941 063 |
| Exc. Hab. le PDD (2) | 8 402 505 | 8 402 505 | 8 402 505 | 8 402 505 |
| Tour. Héb. hors PDD (3) | 10 513 457 | 2 230 039 | 1 160 191 | 1 116 312 |
| Tour. Héb. au PDD (4) | 119 429 939 | 119 429 939 | 65 013 864 | 63 869 182 |
| Dépenses Importance économique (1) + (2) + (3) + (4) | 143 225 739 | 134 003 547 | 78 517 624 | 77 329 063 |
| Dépenses Impact économique (1) + (3) + (4) | 134 823 234 | 125 601 042 | 70 115 119 | 68 926 558 |

Tableau 10 : Dépenses totales selon le type de visiteurs et le type de dépenses considérées dans les autres départements (euros)

(Hyp. 800 000 visiteurs)

| Hyp. 800 000 visiteurs | Dépense totale dans le département concerné | Dépense totale affectable | Dépense totale affectable |
|--|---|---------------------------|---------------------------|
| Catégorie de visiteurs | | (C1) à la CPFL | (C2) à la CPFL |
| Exc. Hab. l'Allier (1) | 162 824 | 162 824 | 162 824 |
| Tour. Héb. dans l'Allier (2) | 3 680 943 | 1 638 823 | 1 261 788 |
| Importance économique (1) + (2) | 3 843 768 | 1 801 647 | 1 424 613 |
| Impact économique (2) | 3 680 943 | 1 638 823 | 1 261 788 |
| Exc. Hab. le Cantal (1) | 65 747 | 65 747 | 65 747 |
| Tour. Héb. dans le Cantal (2) | 2 000 763 | 994 027 | 916 916 |
| Importance économique (1) + (2) | 2 066 510 | 1 059 774 | 982 662 |
| Impact économique (2) | 2 000 763 | 994 027 | 916 916 |
| Exc. Hab. la Haute-Loire (1) | 45 577 | 45 577 | 45 577 |
| Tour. Héb. Dans la Haute-Loire (2) | 813 201 | 380 499 | 447 064 |
| Importance économique (1) + (2) | 858 778 | 426 076 | 492 641 |
| Impact économique (2) | 813 201 | 380 499 | 447 064 |

En termes de **dépense totale**, pour la première des estimations-plancher (**800 000 visiteurs**), l'étude indique que la fréquentation de la CPFL génère en chiffres arrondis, et en Me (millions d'euros) :

134 Me dans le Puy-de-Dôme ;

3,8 Me dans l'Allier ; 2 Me dans le Cantal et 0,9 Me dans la Haute-Loire.

Cependant, ces dépenses ne sont pas toutes attribuables à l'existence de la CPFL, il a donc été estimé un coefficient d'affectation (C1 ou C2) pour chaque visiteur enquêté, comme indiqué dans une section précédente (en utilisant une série de variables de motivation et d'une question de scénario pour identifier les choix de destinations touristiques alternatives si la CPFL n'existaient pas dans le Puy-de-Dôme (comportement de substitution)).

Il a ainsi été possible d'attribuer une proportion des dépenses directes totales dans chaque département. Bien que les limites du scénario et des projections soient reconnues, cela fournit une estimation du montant des dépenses touristiques nouvelles ou conservées attribuables à la CPFL qui n'auraient pas eu lieu autrement.

Ainsi, en termes **d'importance économique**, la fréquentation de la CPFL génère :

78,5 Me (C1) ou **77 Me (C2)** dans le Puy-de-Dôme ;

1,8 Me (C1) ou 1,4 Me (C2) dans l'Allier ; 1 Me (C1) ou 0,99 Me dans le Cantal

et 0,4 Me (C1) ou 0,5 Me (C2) dans la Haute-Loire.

En termes **d'impact économique**, cette fréquentation génère :

70 Me (C1) ou **68,9 Me (C2)** dans le Puy-de-Dôme ;

1,6 Me (C1) ou 1,3 Me (C2) dans l'Allier ; 1 Me (C1) ou 0,9 Me (C2) dans le Cantal

et 0,4 Me (C1) ou 0,45 Me (C2) dans la Haute-Loire.

Présentation simplifiée des résultats selon les deux hypothèses de fréquentation

(le tableau qui suit est celui qui est présenté lors des réunions de restitution en cours et en fin de projet)

| | Hypothèse : 800 000 visiteurs | | Hypothèse : 900 000 visiteurs | |
|---|-------------------------------|----------------------|-------------------------------|----------------------|
| Dépense en millions d'euros | Total | Affectable CPFL (C2) | Total | Affectable CPFL (C2) |
| Dépense partout <i>(importance économique totale)</i> | 143 | 82 | 161 | 92 |
| Dépense dans le 63 <i>(importance économique locale)</i> | 134 | 77 | 150 | 87 |
| Dépense dans le 63 hors dép des résidents 63 <i>(impact économique local)</i> | 126 | 69 | 141 | 78 |

b) Questions d'agrégation

Comme indiqué précédemment, dans le cadre de l'analyse d'impact économique, l'agrégation des impacts économiques des quatre départements *n'est pas équivalente* à l'impact économique de la fréquentation de la CPFL sur la région Auvergne.

En effet, si l'ancienne région Auvergne est l'unité géographique d'analyse, tous les visiteurs qui résident dans les quatre départements deviennent automatiquement des visiteurs locaux du point de vue de la région. Ainsi donc leurs dépenses ne sont pas prises en compte dans le cadre de l'impact économique sur la région. En revanche, cette agrégation peut être appropriée si toutes les dépenses, quel que soit le type de visiteurs, sont prises en compte. C'est le cas d'une analyse de l'importance économique.

c) Simulation du même calcul sur la base d'Unités de Consommation

En section précédente xxx relative aux imputations, il a été indiqué que les calculs de dépenses moyennes dans la voie « principale » ont été réalisés sur la base de « l'effectif-budget total » additionnant adultes et enfants, et qu'une simulation a été faite adoptant la notion « d'Unité de Consommation » (ici selon l'échelle dite d'Oxford). Cette simulation affectait très peu les totaux des dépenses par postes (en partie du fait de la part très minoritaire des imputations faites sur valeurs absentes).

A ce stade du calcul des effets, le même changement de méthode ne peut jouer que si les proportions entre touristes et excursionnistes diffèrent entre « effectifs budgets totaux » et « Unités de consommation » or elles diffèrent très peu, respectivement : 28,1% et 71,9% contre 28,3% et 71,7% pour les calculs en UC. La simulation complète montre ainsi une très faible différence de montant « d'Impact local » de +1,4% (69,9 Me contre 68,9 déjà indiqués).

5. La conversion des dépenses directes en valeur ajoutée, emplois et de revenus.

Comme indiqué précédemment, les dépenses totales peuvent être converties en valeur ajoutée, emplois et de revenus à partir d'un modèle d'impact économique ou d'un ensemble de multiplicateurs. La méthodologie utilisée est celle qui a été construite et testée par Boncoeur et al. (2013) dans le cadre du projet BECO⁵⁶ évaluant les retombées économiques de la fréquentation des sites littoraux et insulaires protégés (îles Chausey et du Mont Saint-Michel) pour obtenir les coefficients d'emploi, de valeur ajoutée et de masse salariale par unité de dépense. Cette méthodologie consiste en fait à effectuer les opérations suivantes :

a) Estimation de l'emploi salarié imputable au tourisme dans chaque département de l'Auvergne

Pour estimer l'emploi salarié imputable au tourisme dans chaque département, les données les suivantes sont utilisées :

- INSEE, recensement de la population : emploi au lieu de travail, niveau départemental.
- INSEE, estimations de l'emploi salarié touristique.

Les données d'emploi au lieu de travail utilisées sont issues du recensement de la population 2017 (Tableau 14). En effet, les résultats statistiques du recensement de la population de l'année N sont mis en ligne au mois de juin de l'année N+3.

Tableau 11 : Emploi total dans les départements en 2017

| Département | Salariés | Non-salariés | Ensemble |
|-------------|----------|--------------|----------|
| Allier | 104 007 | 18 931 | 122 938 |
| Cantal | 44 793 | 12 958 | 57 751 |
| Haute-Loire | 66 165 | 15 239 | 81 404 |
| Puy-de-Dôme | 231 718 | 35 024 | 266 742 |

Source : Insee, RP2017, exploitations principales lieu de travail, géographie au 01/01/2020.

En appliquant le taux d'emploi salarié touristique correspondant à chaque département, on obtient l'emploi salarié touristique de chaque département. Ces proportions (part de l'emploi salarié touristique dans l'emploi salarié total) utilisées, Tableau 15 publié dans : « L'emploi salarié touristique en Auvergne » (SPOT Observatoire Tourisme / INSEE Auvergne, 2011),

⁵⁶ Bountfiles ECONomie

sont issues des Déclarations Annuelles de Données Sociales (DADS⁵⁷) pour l'année 2007. En effet, Baccaïni, Thomas et Khiati, (2006) ont élaboré une méthode d'estimation de l'emploi touristique salarié direct⁵⁸. Exploitant les données issues des DADS, cette méthode se base, d'une part sur la typologie des communes à savoir : bien équipée, moyennement équipée et peu équipée pour le tourisme, et d'autre part sur les types d'activité à savoir : 100 % touristique, fortement touristique, moyennement touristique, faiblement touristique, non touristique. Ainsi, tout ou partie de l'emploi d'un établissement sera considéré comme touristique en fonction de la nature de son activité et du niveau d'équipement touristique de la commune où il est établi.

Tableau 12 : Part et emploi salarié touristique

| Département | Part de l'emploi salarié touristique dans l'emploi salarié total | Emploi salarié touristique |
|-------------|--|----------------------------|
| Allier | 4,40% | 4 576 |
| Cantal | 5,80% | 2 598 |
| Haute-Loire | 3,30% | 2 183 |
| Puy-de-Dôme | 4,50% | 10 427 |

Source : Insee, RP2017, DADS 2007

b) Estimation de l'emploi non salarié imputable au tourisme dans chaque département de l'Auvergne

Hormis l'emploi salarié, on trouve aussi l'emploi non salarié dans le secteur touristique (patron d'hébergement ou de restaurant, commerçant ou artisan, par exemple). Contrairement à l'emploi salarié touristique estimable à partir du DADS, l'emploi non salarié est estimé à partir de l'exploitation des fichiers de l'Agence Centrale des Organismes de Sécurité Sociale (ACOSS) pour les non-salariés non agricoles. Malheureusement, ces données ne sont pas publiées au niveau local par l'Insee. Pour contourner cette difficulté, les **données nationales** relatives à l'emploi non salarié et salarié dans les principales branches directement en contact avec les touristes sont utilisées. Ces branches sont les suivantes : hébergement-restauration⁵⁹,

⁵⁷ Dans ce document, les employeurs fournissent annuellement, entre autres, pour leur établissement les effectifs employés.

⁵⁸ Emploi impliquant un contact direct avec les touristes.

⁵⁹ Hôtellerie, autres hébergements touristiques, restauration

commerce de détail⁶⁰ et autres services⁶¹. *L'hypothèse faite ici, est que la répartition entre l'emploi salarié et non-salarié est la même à l'échelle locale (département) qu'à l'échelle nationale.*

Les sources de données sont à cet effet utilisées :

- Estimation de l'emploi salarié touristique dans chaque département (tableau 15) ;
- INSEE, estimation de l'emploi salarié touristique par type d'activité (exploitation des DADS), années 2007, niveau départemental ;
- INSEE, comptes nationaux 2017 : emploi salarié et non salarié dans les branches d'activité A17.IZ (hébergement-restauration), A88.47 (commerce de détail hors auto-moto) et A10.RU (autres services).

A partir de ces sources de données, les opérations suivantes permettent d'estimer l'emploi non salarié touristique :

- Ventiler l'emploi salarié touristique dans chaque département par type d'activité (utilisation de la structure départementale de l'emploi salarié touristique, Tableau 16) ;
- Calculer le ratio de l'emploi non salarié à l'emploi salarié dans chacune des branches d'activité correspondantes (utilisation des données nationales de branche) ;
- Appliquer ces ratios aux estimations locales d'emploi salarié touristique par type d'activité.

Tableau 13 : Répartition de l'emploi salarié touristique selon les activités

| Type d'activité | Allier | Cantal | Haute-Loire | Puy-de-Dôme | Auvergne | France de province |
|-----------------------------|--------|--------|-------------|-------------|----------|--------------------|
| Hébergement restauration | 44% | 45% | 55% | 50% | 48% | 53% |
| Commerce de détail | 24% | 23% | 25% | 18% | 21% | 18% |
| Autres services | 32% | 32% | 20% | 33% | 31% | 29% |
| Total | 100% | 100% | 100% | 100% | 100% | 100% |

Source : Insee, DADS 2007

⁶⁰ Commerce de détail alimentaire, commerce de détail non alimentaire

⁶¹ Autres activités, activités sportives et récréatives, parcs d'attractions et casinos, thermes et thalassothérapie.

Tableau 14 : Ventilation par activité de l'emploi salarié touristique dans les départements

| Activités | Allier | Cantal | Haute-Loire | Puy-de-Dôme |
|-----------------------------|--------|--------|-------------|-------------|
| Hébergement restauration | 2 032 | 1 164 | 1 205 | 5 172 |
| Commerce de détail | 1 080 | 595 | 550 | 1 825 |
| Autres services | 1 464 | 839 | 428 | 3 430 |
| Total | 4 576 | 2 598 | 2 183 | 10 427 |

Sources : INSEE, RP 2017 et DADS 2007

A l'échelle nationale, on dispose, pour chacune des trois activités ci-dessus, de la ventilation de l'emploi total entre emploi salarié et emploi non salarié :

Tableau 15 : Données nationales sur la structure de l'emploi (en millier) dans les branches « hébergement-restauration », « commerce de détail » et « autres services »

| Branches d'activités Code Libellé | Emplois salariés | Emplois non-salariés | Emploi total | Ratio salariés /non- salariés |
|---------------------------------------|---------------------|-------------------------|-----------------|-------------------------------------|
| A17.IZ Hébergement et restauration | 1 027 | 191 | 1 218 | 0,186 |
| A88.47 Commerce de détail* | 1 824 | 306 | 2 131 | 0,168 |
| A10.RU Autres services | 1 313 | 193 | 1 506 | 0,147 |

* hors automobiles et motocycles. Source : INSEE, comptes nationaux 2017

Pour chaque branche et département, les ratios nationaux calculés sont appliqués aux emplois salariés touristiques pour estimer l'emploi non salarié touristique local.

Tableau 16 : Estimation de l'emploi touristique total dans l'Allier en 2017 : répartition par activité

| Type d'activité | Emploi touristique salarié | | Emploi touristique non salarié | | Emploi Total | touristique Structure |
|--------------------------|----------------------------|-----------|--------------------------------|-----------|--------------|-----------------------|
| | Effectifs | Structure | Effectifs | Structure | | |
| Hébergement restauration | 2 032 | 44% | 378 | 49% | 2 410 | 45% |
| Commerce de détail | 1 080 | 24% | 181 | 23% | 1 261 | 24% |
| Autres services | 1 464 | 32% | 216 | 28% | 1 680 | 31% |
| Total | 4 576 | 100% | 775 | 100% | 5 351 | 100% |

Sources : INSEE, RP 2017, DADS 2007, comptes nationaux 2017

Tableau 17 : Estimation de l'emploi touristique total dans le Cantal en 2017 : répartition par activité

| Type d'activité | Emploi touristique salarié | | Emploi touristique non salarié | | Emploi Total | touristique Structure |
|--------------------------|----------------------------|-----------|--------------------------------|-----------|--------------|-----------------------|
| | Effectifs | Structure | Effectifs | Structure | | |
| Hébergement restauration | 1 164 | 45% | 216 | 49% | 1 380 | 45% |
| Commerce de détail | 595 | 23% | 100 | 23% | 695 | 23% |
| Autres services | 839 | 32% | 124 | 28% | 963 | 32% |
| Total | 2 598 | 100% | 440 | 100% | 3 038 | 100% |

Sources : INSEE, RP 2017, DADS 2007, comptes nationaux 2017

Tableau 18 : Estimation de l'emploi touristique total dans la Haute-Loire en 2017 : répartition par activité

| Type d'activité | Emploi salarié | touristique | Emploi non salarié | touristique | Emploi Total | touristique |
|--------------------------|----------------|-------------|--------------------|-------------|--------------|-------------|
| | Effectifs | Structure | Effectifs | Structure | Effectifs | Structure |
| Hébergement restauration | 1 205 | 55% | 224 | 59% | 1 429 | 56% |
| Commerce de détail | 550 | 25% | 92 | 24% | 642 | 25% |
| Autres services | 428 | 20% | 63 | 17% | 491 | 19% |
| Total | 2 183 | 100% | 379 | 100% | 2 562 | 100% |

Sources : INSEE, RP 2017, DADS 2007, comptes nationaux 2017

Tableau 19 : Estimation de l'emploi touristique total dans la Puy-de-Dôme en 2017 : répartition par activité

| | Emploi salarié | touristique | Emploi non salarié | touristique | Emploi Total | touristique |
|--------------------------|----------------|-------------|--------------------|-------------|--------------|-------------|
| | Effectifs | Structure | Effectifs | Structure | Effectifs | Structure |
| Hébergement restauration | 5 172 | 50% | 962 | 54% | 6 134 | 50% |
| Commerce de détail | 1 825 | 18% | 307 | 17% | 2 132 | 17% |
| Autres services | 3 430 | 33% | 505 | 28% | 3 935 | 32% |
| Total | 10 427 | 100% | 1 774 | 100% | 12 201 | 100% |

Sources : INSEE, RP 2017, DADS 2007, comptes nationaux 2017

Le tableau ci-dessous récapitule l'emploi touristique dans les quatre départements de l'ancienne région Auvergne

Tableau 20 : Estimation de l'emploi touristique total par département en 2017

| Département | Emploi touristique salariés | Emploi touristique non-salariés | Emploi Total Effectifs | Structure en Auvergne | % touristique salariés dans l'emploi touristique total | Proportion d'emplois touristiques dans l'emploi total (en %) |
|-------------|-----------------------------|---------------------------------|------------------------|-----------------------|--|--|
| Allier | 4 576 | 775 | 5 351 | 23% | 86% | 4% |
| Cantal | 2 598 | 440 | 3 038 | 13% | 86% | 5% |
| Haute-Loire | 2 183 | 379 | 2 562 | 11% | 85% | 3% |
| Puy-de-Dôme | 10 427 | 1 774 | 12 201 | 53% | 85% | 5% |
| Auvergne | 19 784 | 3 368 | 23 152 | 100% | 85% | |

Sources : INSEE, RP 2017, DADS 2007, comptes nationaux 2017

c) Estimation des coefficients d'emploi, de valeur ajoutée et de masse salariale adaptés au contexte touristique local

La conversion de la dépense totale des visiteurs en valeur ajoutée, emplois et de revenus, passe par l'estimation de coefficients à partir des données correspondantes aux branches d'activité directement concernées par cette dépense. Ces branches sont : hébergement-restauration (branche A17.IZ), commerce de détail (branche A88.47), autres services (branche A10.RU). Malheureusement, ces données sont généralement non disponibles au niveau local (département dans notre contexte). Pour contourner cette difficulté, les données nationales sont donc adaptées au contexte socioéconomique local. Ainsi donc, les comptes nationaux permettent d'avoir les données liées à l'emploi total, à la production, à la valeur ajoutée HT et à la rémunération des salariés ou masse salariale (salaires bruts et cotisations sociales patronales) de chaque branche. Les ventes TTC des branches hébergement-restauration et autres services sont obtenues en ajoutant à la production de la branche le montant des impôts sur les produits (notamment la TVA). Ainsi dans ce travail, les taux de TVA retenus sont de 10% et 5,5% respectivement pour la branche hébergement-restauration et autres services. De son côté, les comptes du commerce fournissent le montant des ventes TTC de la branche commerce de détail. L'estimation des coefficients suit les étapes suivantes :

- Calcul des ratios nationaux ventes TTC, valeur ajoutée HT et masse salariale par unité d'emploi des branches hébergement-restauration (branche A17.IZ), commerce de détail (branche A88.47), autres services (branche A10.RU) ;
- Adaptation de ces ratios à l'économie départementale et à l'activité touristique ;
- Pondération des ratios corrigés de branche à l'aide de la structure de l'emploi touristique (Tableaux 19, 20, 21, 22) ;
- Calcul de coefficients d'emploi, de valeur ajoutée et de masse salariale par unité de dépense touristique.

Tableau 21 : Données nationales sur les branches « hébergement-restauration », « Commerce de détail » et « autres services »

| | Emploi total (milliers) | Ventes TTC (milliards d'euros) | Valeur ajoutée (milliards d'euros) | Masse salariale (milliards d'euros) |
|------------------------------------|-------------------------|--------------------------------|------------------------------------|-------------------------------------|
| A17.IZ Hébergement et restauration | 1 217,66 | 118,97 | 55,48 | 35,33 |
| A88.47 Commerce de détail | 2 130,51 | 506,30 | 85,42 | 57,11 |
| A10.RU Autres services | 1 506,26 | 107,56 | 60,72 | 45,66 |

Source : INSEE, comptes nationaux et comptes du commerce 2017.

Calcul des ratios nationaux ventes TTC, valeur ajoutée HT et masse salariale par unité d'emploi de chaque branche.

Tableau 22 : Ratios nationaux de branche (euros par emploi, 2017)

| | Ventes TTC / Emploi | Valeur ajoutée / Emploi | Masse salariale / Emploi |
|------------------------------------|---------------------|-------------------------|--------------------------|
| A17.IZ Hébergement et restauration | 97 701 | 45 566 | 29 018 |
| A88.47 Commerce de détail | 237 642 | 40 095 | 26 807 |
| A10.RU Autres services | 71 410 | 40 313 | 30 311 |

Source : INSEE, comptes nationaux et comptes du commerce 2017.

L'adaptation de ces ratios nationaux à l'économie départementale passe par deux types de corrections.

a) La première correction

Cette première correction permet de prendre en compte l'écart entre les prix et revenus à l'échelle départementale et à l'échelle nationale. Cet écart est appréhendé à travers le salaire horaire moyen. Ce dernier dans l'Allier, le Cantal, la Haute-Loire et le Puy-de-Dôme représentait respectivement, 84% (0,84), 79% (0,79), 82% (0,82) et 93% (0,93) dans l'Allier, le Cantal, la Haute-Loire, le Puy-de-Dôme du niveau national en 2017 ([Annexe 1](#)). Les ratios sont donc multipliés par ces proportions.

b) La seconde correction

Cette seconde correction permet de prendre en compte l'importance de l'emploi saisonnier et à temps partiel dans le secteur touristique. Ce phénomène est saisi à travers l'écart entre les ratios de l'effectif mesuré en équivalent temps plein (ETP) à l'effectif annuel moyen dans l'emploi salarié touristique d'une part et dans l'emploi salarié au sein de l'ensemble de l'économie d'autre part. Le ratio ETP sur emploi salarié non touristique n'étant pas disponible au niveau départemental, c'est celui de l'ancienne région Auvergne qui est utilisé. Les résultats obtenus lors de la première correction sont multipliés par les ratios des Tableaux 26, 27, 28 et 29 afin de les adapter au contexte touristique.

Tableau 23 : Effet de l'emploi saisonnier et à temps partiel (Allier)

| | Emploi salarié touristique [1] | Emploi salarié non touristique | Emploi salarié total [2] | Ratio [1] / [2] |
|--|--------------------------------|--------------------------------|--------------------------|-----------------|
| Effectif annuel moyen | 4 514 | 98 077 | 102 591 | 4,40% |
| Equivalent temps plein | 3 431 | 80 423 | 83 854 | 4,09% |
| Equivalent temps plein / effectif annuel moyen | 0,76 | 0,82 | 0,82 | 0,93 |

Source : Insee, DADS 2007

Tableau 24 : Effet de l'emploi saisonnier et à temps partiel (Cantal)

| | Emploi salarié touristique [1] | Emploi salarié non touristique | Emploi salarié total [2] | Ratio [1] / [2] |
|--|--------------------------------|--------------------------------|--------------------------|-----------------|
| Effectif annuel moyen | 2 452 | 39 824 | 42 276 | 5,80% |
| Equivalent temps plein | 1 913 | 32 656 | 34 568 | 5,53% |
| Equivalent temps plein / effectif annuel moyen | 0,78 | 0,82 | 0,82 | 0,95 |

Source : Insee, DADS 2007

Tableau 25 : Effet de l'emploi saisonnier et à temps partiel (Haute-Loire)

| | Emploi salarié touristique [1] | Emploi salarié non touristique | Emploi salarié total [2] | Ratio [1] / [2] |
|--|--------------------------------|--------------------------------|--------------------------|-----------------|
| Effectif annuel moyen | 2 076 | 60 833 | 62 909 | 3,30% |
| Equivalent temps plein | 1 515 | 49 883 | 51 399 | 2,95% |
| Equivalent temps plein / effectif annuel moyen | 0,73 | 0,82 | 0,82 | 0,89 |

Source : Insee, DADS 2007

Tableau 26 : Effet de l'emploi saisonnier et à temps partiel (Puy-de-Dôme)

| | Emploi salarié touristique [1] | Emploi salarié non touristique | Emploi salarié total [2] | Ratio [1] / [2] |
|--|--------------------------------|--------------------------------|--------------------------|-----------------|
| Effectif annuel moyen | 9 732 | 206 535 | 216 267 | 4,50% |
| Equivalent temps plein | 7 299 | 169 358 | 176 657 | 4,13% |
| Equivalent temps plein / effectif annuel moyen | 0,75 | 0,82 | 0,82 | 0,92 |

Source : Insee, DADS 2007

Tableau 27 : Récapitulatifs des deux corrections

| Allier | Cantal | Haute-Loire | Puy-de-Dôme |
|--------|--------|-------------|-------------|
|--------|--------|-------------|-------------|

| | | | | | | | | | |
|--|---|----------------------------|---|----------------------------|---|----------------------------|---|----------------------------|---|
| Coefficient d'adaptation l'économie locale | à | $0,84*0,93$ 0,78 | = | $0,79*0,95$ 0,75 | = | $0,82*0,89$ 0,73 | = | $0,93*0,92$ 0,86 | = |
|--|---|----------------------------|---|----------------------------|---|----------------------------|---|----------------------------|---|

Source : Insee, DADS 2007, Salaire net horaire moyen 2017

Les ratios nationaux de branche adaptés à l'économie locale sont ensuite pondérés à l'aide de la structure de l'emploi touristique (Tableaux 19, 20, 21, 22) pour obtenir des ratios représentatifs de la consommation touristique. Cette opération permet d'obtenir des ratios moyens de ventes, de valeur ajoutée et de masse salariale par emploi (toutes activités confondues) cohérents avec la consommation des touristes présents dans chaque zone d'étude.

Tableau 28 : Ratios de branche corrigés (euros par emploi, 2017, Allier)

| | Ventes TTC / Emploi | Valeur ajoutée / Emploi | Masse salariale / Emploi |
|------------------------------------|------------------------|----------------------------|-----------------------------|
| A17.IZ Hébergement et restauration | 75 926 | 35 410 | 22 551 |
| A88.47 Commerce de détail | 184 677 | 31 159 | 20 832 |
| A10.RU Autres services | 55 494 | 31 328 | 23 555 |

Sources : INSEE, élaboration propre

Tableau 29 : Ratios de branche corrigés pondérés (Allier)

| | Coefficients de pondération* | Ventes TTC / Emploi | Valeur ajoutée / Emploi | Masse salariale / Emploi |
|------------------------------------|---------------------------------|------------------------|----------------------------|-----------------------------|
| A17.IZ Hébergement et restauration | 45% | 34 196 | 15 948 | 10 156 |
| A88.47 Commerce de détail | 24% | 43 520 | 7 343 | 4 909 |
| A10.RU Autres services | 31% | 17 423 | 9 836 | 7 395 |
| Total | 100% | 95 139 | 33 127 | 22 461 |

* Structure locale de l'emploi touristique. Sources : INSEE, élaboration propre

Tableau 30 : Ratios de branche corrigés (euros par emploi, 2017, Cantal)

| | Ventes TTC / Emploi | Valeur ajoutée / Emploi | Masse salariale / Emploi |
|------------------------------------|------------------------|----------------------------|-----------------------------|
| A17.IZ Hébergement et restauration | 73 359 | 34 213 | 21 788 |
| A88.47 Commerce de détail | 178 435 | 30 106 | 20 128 |
| A10.RU Autres services | 53 619 | 30 269 | 22 759 |

Sources : INSEE, élaboration propre

Tableau 31 : Ratios de branche corrigés pondérés (Cantal)

| | Coefficients de pondération* | Ventes TTC / Emploi | Valeur ajoutée / Emploi | Masse salariale / Emploi |
|------------------------------------|------------------------------|------------------------|----------------------------|-----------------------------|
| A17.IZ Hébergement et restauration | 45% | 33 323 | 15 541 | 9 897 |
| A88.47 Commerce de détail | 23% | 40 820 | 6 887 | 4 605 |
| A10.RU Autres services | 32% | 16 996 | 9 595 | 7 214 |
| Total | 100% | 91 140 | 32 023 | 21 716 |

* Structure locale de l'emploi touristique. Sources : INSEE, élaboration propre

Tableau 32 : Ratios de branche corrigés (euros par emploi, 2017, Haute-Loire)

| Haute-Loire | | Ventes TTC / Emploi | Valeur ajoutée / Emploi | Masse salariale / Emploi |
|------------------------------------|----|------------------------|----------------------------|-----------------------------|
| A17.IZ Hébergement et restauration | et | 71 617 | 33 401 | 21 271 |
| A88.47 Commerce de détail | | 174 198 | 29 391 | 19 650 |
| A10.RU Autres services | | 52 345 | 29 550 | 22 219 |

Sources : INSEE, élaboration propre

Tableau 33 : Ratios de branche corrigés pondérés (Haute-Loire)

| | | Coefficients de pondération* | Ventes TTC / Emploi | Valeur ajoutée / Emploi | Masse salariale / Emploi |
|------------------------------------|----|------------------------------|------------------------|----------------------------|-----------------------------|
| A17.IZ Hébergement et restauration | et | 56% | 39 946 | 18 630 | 11 864 |
| A88.47 Commerce de détail | | 25% | 43 651 | 7 365 | 4 924 |
| A10.RU Autres services | | 19% | 10 032 | 5 663 | 4 258 |
| Total | | 100% | 93 629 | 31 658 | 21 046 |

* Structure locale de l'emploi touristique. Sources : INSEE, élaboration propre

Tableau 34 : Ratios de branche corrigés (euros par emploi, 2017, Puy-de-Dôme)

| Puy-de-Dôme | Ventes TTC / Emploi | Valeur ajoutée / Emploi | Masse salariale / Emploi |
|------------------------------------|------------------------|----------------------------|-----------------------------|
| A17.IZ Hébergement et restauration | 83 838 | 39 101 | 24 901 |
| A88.47 Commerce de détail | 203 924 | 34 406 | 23 003 |
| A10.RU Autres services | 61 278 | 34 593 | 26 010 |

Sources : INSEE, élaboration propre

Tableau 35 : Ratios de branche corrigés pondérés (Puy-de-Dôme)

| | Coefficients de pondération* | Ventes TTC / Emploi | Valeur ajoutée / Emploi | Masse salariale / Emploi |
|------------------------------------|------------------------------|------------------------|----------------------------|-----------------------------|
| A17.IZ Hébergement et restauration | 50% | 42 149 | 19 658 | 12 519 |
| A88.47 Commerce de détail | 17% | 35 634 | 6 012 | 4 020 |
| A10.RU Autres services | 32% | 19 763 | 11 157 | 8 389 |
| Total | 100% | 97 546 | 36 827 | 24 927 |

* Structure locale de l'emploi touristique. Sources : INSEE, élaboration propre

Obtention des coefficients d'emploi, de valeur ajoutée et de masse salariale par unité de dépense à partir des tableaux précédents :

a) Coefficient d'emploi = inverse du ratio ventes TTC / emploi.

Afin d'améliorer la lisibilité, ce coefficient est exprimé en termes d'emplois directs générés par un million d'euros de dépense touristique. Aussi, il est exprimé en termes d'ETP annuels. Pour obtenir les ETP annuels à partir des emplois annuels, il faut multiplier le nombre d'emplois annuel par le ratio du nombre d'ETP salariés au nombre d'emplois salariés touristiques. Ces ratios pour les départements Allier, Cantal, Haute-Loire et Puy-de-Dôme se trouvent respectivement dans les Tableaux 26, 27, 28 et 29. Par ailleurs, à partir d'hypothèses concernant les emplois non-salariés, le coefficient d'emploi en termes d'ETP est présenté sous la forme d'un intervalle. Le calcul de la borne inférieure de cet intervalle est basée sur l'hypothèse que le temps partiel affecte de la même façon les emplois salariés et non-salariés touristiques. De ce fait, les ratios sont appliqués à l'ensemble de l'emploi touristique. Le calcul de la borne supérieure repose, pour sa part, sur l'hypothèse que le temps partiel n'affecte que les emplois salariés, tous les emplois non-salariés étant supposés à temps plein. Par conséquent, on applique les ratios aux seuls emplois salariés touristiques.

b) Coefficient de valeur ajoutée = (valeur ajoutée / emploi) / (ventes TTC / emploi)

c) Coefficient de masse salariale = (masse salariale / emploi) / (ventes TTC / emploi)

Tableau 36 : Effets directs de la dépense touristique : coefficients d'emploi, de valeur ajoutée et de masse salariale

| | Nombre d'emploi pour 1 million d'euros de dépenses TTC | | Valeur ajoutée / dépenses TTC | Masse salariale / dépenses TTC |
|-------------|--|------------------------|-------------------------------|--------------------------------|
| | Effectif annuel moyen | Equivalent temps plein | | |
| Allier | 10,51 | 7,99 - 8,35 | 0,35 | 0,24 |
| Cantal | 10,97 | 8,56 - 8,91 | 0,35 | 0,24 |
| Haute-Loire | 10,68 | 7,80 - 8,22 | 0,34 | 0,22 |
| Puy-de-Dôme | 10,25 | 7,69 - 8,06 | 0,38 | 0,26 |

Sources : INSEE, élaboration propre

d) Estimation de la valeur ajoutée, de la masse salariale et des emplois directement générés par les dépenses des visiteurs

Ainsi 1 Million d'euros de dépenses des visiteurs créent :

- 10,25 emplois directs⁶²
- Entre 7,69 et 8,06 équivalents temps plein (selon nature des emplois non-salariés)
- 0,38 millions d'euros en valeur ajoutée
- Dont 0,26 en masse salariale

Tableau 37 : Effets directs des dépenses des visiteurs de la CPFL sur l'économie des départements

| Hyp. 800 000 visiteurs | Type de dépense | Nombre d'emplois | Equivalents temps plein intervalle | | Valeur ajoutée | Masse salariale |
|------------------------|-------------------------------|------------------|------------------------------------|-------|----------------|-----------------|
| PDD | Totale | 1 288 | 966 | 1 013 | 47 418 006 | 32 096 073 |
| | <i>Affectable (C1) à CPFL</i> | 719 | 539 | 565 | 26 470 470 | 17 917 205 |
| | <i>Affectable (C2) à CPFL</i> | 707 | 530 | 556 | 26 021 783 | 17 613 500 |
| Allier | Totale | 39 | 29 | 31 | 1 281 650 | 869 003 |
| | <i>Affectable (C1) à CPFL</i> | 17 | 13 | 14 | 570 627 | 386 905 |
| | <i>Affectable (C2) à CPFL</i> | 13 | 10 | 11 | 439 346 | 297 892 |
| Cantal | Totale | 22 | 17 | 18 | 703 001 | 476 731 |
| | <i>Affectable (C1) à CPFL</i> | 11 | 9 | 9 | 349 268 | 236 851 |
| | <i>Affectable (C2) à CPFL</i> | 10 | 8 | 8 | 322 173 | 218 478 |
| Haute-Loire | Totale | 9 | 6 | 7 | 274 962 | 182 796 |
| | <i>Affectable (C1) à CPFL</i> | 4 | 3 | 3 | 128 656 | 85 531 |
| | <i>Affectable (C2) à CPFL</i> | 5 | 3 | 4 | 151 163 | 100 493 |

En résumé, concernant le seul département du PDD

| | Dép. en 63 attribuables CPFL | Nombre d'emplois directs | Equivalent temps plein | Valeur ajoutée | Masse salariale |
|------------------------------------|------------------------------|--------------------------|------------------------|----------------|-----------------|
| | Me | | | Me | Me |
| Coefficients | 1 | 10,25 | [7,69 - 8,06] | 0,38 | 0,26 |
| Hypothèse 800 000 visiteurs | 69 | 707 | [530 - 556] | 26 | 18 |
| Hypothèse 900 000 visiteurs | 78 | 795 | [596 - 625] | 29 | 20 |

⁶² Ce résultat est à rapprocher du coefficient indiqué par (Min-env, 2017) $35\% / 32\ 672 = 10,7$

B. Estimation des effets induits par les effets directs

1. Méthodologie

a) Rappel des limites

Dans cette partie du projet Valo-sites :

- les effets directs sont supposés proportionnels aux flux monétaires.
- les seuls effets induits estimés sont *ceux qui découlent des effets directs*.

Or :

- les évolutions de fréquentation peuvent créer des **modifications de l'offre**, de façon plus complexe (effets de seuil, etc.)
- les effets **indirects** de la fréquentation (voir Tâches 2 et 4.2) *créent eux-mêmes des effets induits*,

On peut ainsi envisager d'autres effets (de plus long terme), dans une autre tâche:

- Directs, par modification de l'offre
- Induits, supplémentaires, par le jeu des effets indirects

b) Prise en compte du caractère plus ou moins résidentiel des zones d'emploi

La méthodologie utilisée pour estimer les effets induits se base sur la théorie la base économique. Les données de l'INSEE relatives à l'emploi dans la sphère présentielle et la sphère productive montrent une sorte de proportionnalité⁶³ entre les emplois des deux sphères économiques. En effet, comme le montre la figure 3 portant sur les données des départements semblables à nos départements d'étude (moins de 700 000 résidents), l'emploi dans la sphère présentielle, dans chaque département, est presque corrélé ($R^2 = 0,8612$) à l'emploi dans la sphère productive. Cependant, cette corrélation est plus forte ($R^2 = 0,9487$) entre l'emploi dans la sphère présentielle et la population résidente (figure 4). C'est dû au fait que les besoins des résidents actifs et inactifs (en particulier des retraités) sont pourvus par les emplois présentiels. Ce qui montre le caractère résidentiel des départements.

⁶³ Cette proportionnalité peut être vraie globalement mais pas dans le principe : certains départements peuvent avoir des bases productives très faibles ou très fortes mais des bases présentielles sans aucun lien (tourisme ou attraction de retraités). Les départements auvergnats comme Allier et le Cantal semblent dans ce cas de figure.

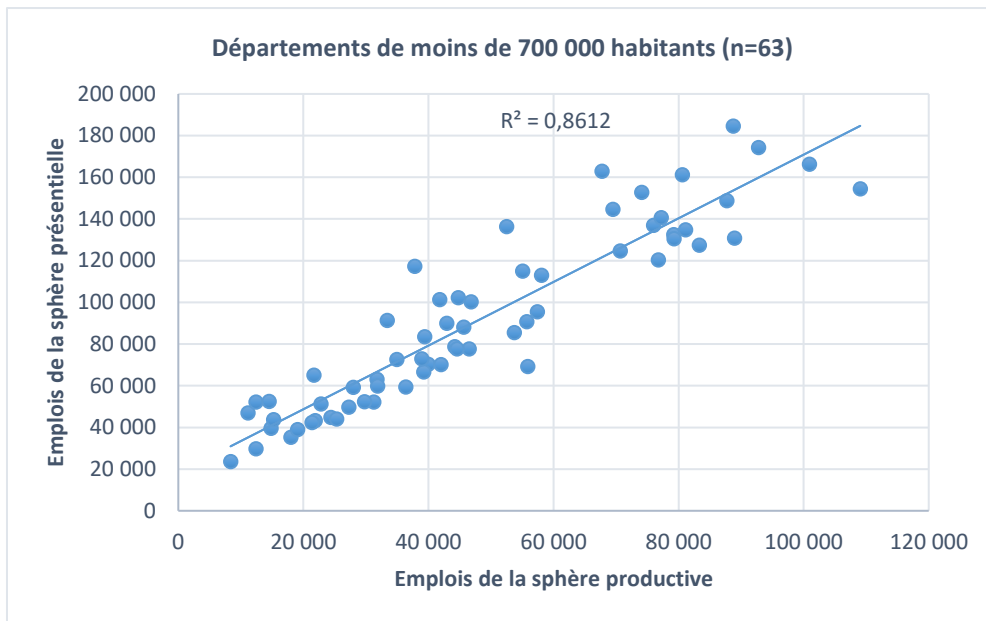


Figure 4 : Corrélation entre emplois productifs et emplois présentiels

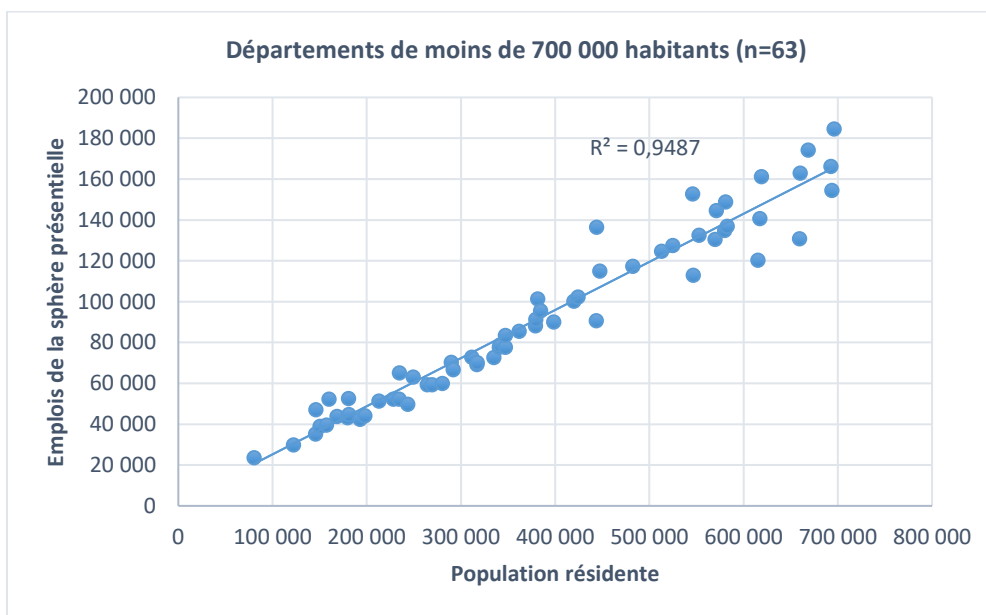


Figure 5 : Corrélation entre population résidente et emplois dans la sphère présentielle

Notons pour chaque département :

P = la population résidente

$E_{Prés}$ = les emplois de la sphère présentielle

E_{Prod} = les emplois de la sphère productive

r = le reste de la population

$$R = E_{Prod} + r$$

- a) **Première hypothèse** : Corrélation entre les emplois dans la sphère présentielle à ceux dans la sphère productive. On a donc :

$$E_{Prés} = a E_{Prod} \quad (1)$$

On peut noter donc qu'un emploi dans la sphère productive induit a emplois dans la sphère présentielle.

- b) **Seconde hypothèse** : Corrélation entre les emplois dans la sphère présentielle à la population résidente.

Cette hypothèse permet de tenir compte du fait que les besoins des résidents actifs et inactifs sont pourvus par les emplois présentiels. On a alors :

$$P = E_{Prés} + R \quad (2)$$

L'hypothèse de proportionnalité entre population résidente et emplois présentiels s'écrit :

$$E_{Prés} = b \times P \quad (3)$$

où b est une constante que l'on peut qualifier de « coefficient d'induction direct ». Les équations (2) et (3) donnent :

$$P = R \times (1 - b)^{-1} \quad (4)$$

où $(1 - b)^{-1}$ est un paramètre jouant un rôle analogue au multiplicateur d'investissement keynésien. Pour obtenir le nombre total d'emplois de la sphère présentielle en fonction de l'effectif du reste de la population de chaque département, on rapproche les équations (3) et (4). On obtient alors :

$$E_{Prés} = R \times b \times (1 - b)^{-1} \quad (5)$$

où le nombre : $c = b \times (1 - b)^{-1}$ peut être qualifié de « coefficient d'induction total ». Il intègre en effet la boucle de rétroaction résultant du fait que les personnes employées dans la sphère présentielle ont elles-mêmes recours aux services fournis par cette sphère. Etant donné que les personnes employées dans la sphère productive forment un sous-ensemble de R , on peut, écrire en première approximation qu'un emploi dans la sphère productive induit donc c emplois dans la sphère présentielle. Cependant, il faut noter que les personnes ayant un emploi peuvent avoir à leur charge des enfants ou des personnes d'âge actif sans emploi. La prise en compte des personnes à charge conduit alors à estimer le nombre moyen d'emplois présentiels induit par un emploi « de base » dans le département :

$$n \times c = n \times b \times (1 - b)^{-1} \quad (6)$$

où n est le nombre moyen de personnes de moins de 60 ans en charge par actif employé dans chaque département.

Par ailleurs, une adaptation des emplois des deux sphères économiques est nécessaire pour le calcul des effets induits. En effet selon la définition INSEE (<https://www.insee.fr/fr/metadonnees/definition/c1886>) :

« Les activités présentielles sont les activités mises en œuvre localement pour la production de biens et de services visant la satisfaction des besoins de personnes présentes dans la zone, qu'elles soient résidentes ou touristes. Les activités productives sont déterminées par différence. Il s'agit des activités qui produisent des biens majoritairement consommés hors de la zone et des activités de services tournées principalement vers les entreprises de cette sphère ».

Ainsi donc, l'INSEE intègre les emplois touristiques aux emplois de la sphère présentielle. Or, les emplois liés à la fréquentation touristique d'une zone sont, fonctionnellement, des emplois basiques dans le contexte de la théorie de la base économique. Puisque les revenus finançant les activités touristiques viennent en majorité de l'extérieur de la zone considérée. Pour cette raison, les emplois touristiques générés par les dépenses de visiteurs doivent être retranchés des emplois de la sphère présentielle et ajoutés à ceux de la sphère productive.

2. Résultats

a) Données

Tableau 38 : Répartition de l'emploi entre sphères productive et présentielle en 2017

| Départements | Allier | Cantal | Haute-Loire | Puy-de-Dôme |
|---|---------|--------|-------------|-------------|
| Sphère productive (INSEE) (1) | 39 365 | 19 068 | 29 746 | 92 748 |
| Sphère présentielle (INSEE) (2) | 83 495 | 38 936 | 52 324 | 174 257 |
| Total (1) + (2) | 122 859 | 58 004 | 82 070 | 267 005 |
| Emploi lié à la fréquentation touristique (3) | 5 351 | 3 038 | 2 562 | 12 201 |
| Sphère productive corrigée (1) + (3) | 44 716 | 22 106 | 32 308 | 104 949 |
| Sphère présentielle corrigée (2) – (3) | 78 144 | 35 898 | 49 762 | 162 056 |

Sources : (1) et (2) INSEE, RP 2017 (exploitation complémentaire), emploi au lieu de travail ; (3) Elaboration propre. Note : il existe de légères différences entre les volumes d'emploi résultant de l'exploitation complémentaire du RP et celles qui résultent de son exploitation principale.

Les données directement nécessaires au calcul des coefficients d'induction sont regroupées dans le tableau ci-dessous

Tableau 39 : Données pour le calcul des coefficients d'induction dans les départements

| Départements | Allier | Cantal | Haute-Loire | Puy-de-Dôme |
|--|---------|---------|-------------|-------------|
| Population totale | 337 988 | 145 143 | 227 283 | 653 742 |
| dont population âgée de moins de 60 ans | 222 720 | 94 114 | 157 580 | 472 622 |
| Emploi total | 122 859 | 58 004 | 82 070 | 267 005 |
| dont emploi dans la sphère présentielle (corrigée du tourisme) | 78 144 | 35 898 | 49 762 | 162 056 |

Source : INSEE, RP 2017 + élaboration propre pour la correction touristique.

| | Allier | Cantal | Haute-Loire | Puy-de-Dôme |
|--|--------|--------|-------------|-------------|
| (1) $b = E_{Prés} / P$ | 23,12% | 24,73% | 21,89% | 24,79% |
| (2) $c = b \times (1 - b)^{-1}$ | 30,07% | 32,86% | 28,03% | 32,96% |
| (3) $n =$ (Population âgée de moins de 60 ans)/ (Emploi total) | 1,81 | 1,62 | 1,92 | 1,77 |
| (4) $b \times c$ | 54,52% | 53,32% | 53,82% | 58,34% |

b) Résultat final

L'effet de la fréquentation de la CPFL, dans le département du Puy de Dôme, pour une hypothèse de 800 000 visiteurs et selon le coefficient d'affectation C2, s'élève à :

1 119 emplois dont 707 emplois directs et 412 emplois induits.

Pour une hypothèse de 900 000 visiteurs les résultats augmentent en proportion :

1 259 emplois dont 795 emplois directs et 464 emplois induits.

(On rappelle que cette étude ne calcule pas les emplois indirects qui s'ajouteraient aux précédents, ni d'autres emplois induits à partir des emplois indirects.)

Récapitulatif selon les deux hypothèses de fréquentation pour le département du PDD

| Département PDD | Type de dépense | Emplois directs | Emplois induits | Ensemble |
|------------------------|------------------------|-----------------|-----------------|----------|
| Hyp. 800 000 visiteurs | Totale | 1 495 | 872 | 2 366 |
| | Affectable (C2) à CPFL | 707 | 412 | 1 119 |
| Hyp. 900 000 visiteurs | Totale | 1 681 | 981 | 2 662 |
| | Affectable (C2) à CPFL | 795 | 464 | 1 259 |

Tableau 40 : Effets directs et induits des dépenses des visiteurs de la CPFL sur l'emploi dans les départements

| Hyp. 800 000 visiteurs | Type de dépense | Emplois directs | Emplois induits | Ensemble |
|------------------------|------------------------|-----------------|-----------------|----------|
| PDD | Totale | 1 495 | 872 | 2 366 |
| | Affectable (C1) à CPFL | 719 | 419 | 1 138 |
| | Affectable (C2) à CPFL | 707 | 412 | 1 119 |
| Allier | Totale | 33 | 18 | 51 |
| | Affectable (C1) à CPFL | 17 | 9 | 27 |
| | Affectable (C2) à CPFL | 13 | 7 | 20 |
| Cantal | Totale | 23 | 12 | 35 |
| | Affectable (C1) à CPFL | 11 | 6 | 17 |
| | Affectable (C2) à CPFL | 10 | 5 | 15 |
| Haute-Loire | Totale | 19 | 10 | 29 |
| | Affectable (C1) à CPFL | 4 | 2 | 6 |
| | Affectable (C2) à CPFL | 5 | 3 | 7 |

VI. Bibliographie

Agnarsson S., Radford A. et Riddington G. (2008), « Economic Impact of Angling in Scotland and Iceland », in *Global Challenges in Recreational Fisheries*, John Wiley & Sons, Ltd, pp. 188-201.

Ardilly P. (2006), *Les techniques de sondage*, Editions TECHNIP, 700 p.

Ardilly P. et Le Blanc D. (2001), « Échantillonnage et pondération d'une enquête auprès de personnes sans domicile : un exemple français », vol. 27, n°1, pp. 117-127.

Baccaïni B., Thomas G. et Khiati A. (2006), *L'emploi salarié dans le tourisme : une nouvelle estimation*, Insee (Insee Première), 4 p.

Ballantyne R., Brown R., Pegg S. et Scott N. (2008), *Valuing tourism spend arising from visitation to Queensland national parks*, [Gold Coast, Qld.] : Sustainable Tourism CRC.

Bangsund D.A. et Leistritz F.L. (2003), « Resident And Nonresident Hunter And Angler Expenditures, Characteristics, And Economic Effects, North Dakota, 2001-2002 », *Agribusiness & Applied Economics Report*, n°23549, North Dakota State University, Department of Agribusiness and Applied Economics.

Barbu A. (2010), « Methods of sample adjustment. Improving the quality of survey results », *Revista Română de Statistică*, vol. 58.

Beckman E., Kumar A. et Kim Y.-K. (2013), « The Impact of Brand Experience on Downtown Success », *Journal of Travel Research*, vol. 52, n°5, pp. 1-13.

Beckman E. et Traynor M. (2018), « Utilizing trade market analysis to identify the economic impact of a multiday special event in Miami Beach, Florida », *Tourism Economics*, vol. 25, n°2, pp. 1-21.

Bergstrom J.C., Cordell H.K., Ashley G.A. et Watson A.E. (1990), « Economic Impacts of Recreational Spending on Rural Areas: A Case Study », *Economic Development Quarterly*, vol. 4, n°1, pp. 29-39.

Bergstrom J., Cordell H., Watson A.E. et Ashley G.A. (1990), « Economic Impact of State Parks on State Economies in the South », *Southern Journal of Agricultural Economics*, vol. 22, n°2, pp. 69-77.

Boncoeur J., Alban F., Brigand L., Coic F., Guégan C., Guingot M., Le Berre S., Le Corre N., Levi L., Meunier M., Pagès C., Peuziat I. et Voltaire L. (2013), « Programme LITEAU III. Projet de recherche BECO. Evaluation et suivi des effets économiques de la fréquentation des sites littoraux et insulaires protégés : application aux îles Chausey et au Mont Saint-Michel. Rapport final. », Rapport final, *Publications électroniques Amure*, n°Série Rapports R-33-2013.

Bonn M.A. et Harrington J. (2008), « A Comparison of Three Economic Impact Models for Applied Hospitality and Tourism Research », *Tourism Economics*, vol. 14, n°4, pp. 769-789.

Bowker J.M., Bergstrom J.C. et Gill J. (2007), « Estimating the Economic Value and Impacts of Recreational Trails: A Case Study of the Virginia Creeper Rail Trail », *Tourism Economics*, vol. 13, n°2, pp. 241-260.

Brahic É. et Terreaux J.-P. (2010), « Estimer la valeur économique de la biodiversité en forêt, difficultés et méthodes », *Sciences Eaux & Territoires*, vol. Numéro 3, n°3, pp. 16-19.

Brilhault G. et Caron N. (2004), « Correction de la non-réponse totale : par imputation ou par repondération ? », Insee document de travail, n°E2004/01, Paris, INSEE.

Brousse C., Guiot de la Rochère B. et Massé E. (2006), « L'enquête sans-domicile 2001 », *Insee Méthodes*, n°116, Paris, Insee.

Buultjens J. et Luckie K. (2004), « Economic impact of selected national parks in north-eastern New South Wales ».

Carlino G.A. et Mills E.S. (1987), « The determinants of county growth », *Journal of Regional Science*, vol. 27, n°1, pp. 39-54.

Carlsen J. (1997), « Economic Evaluation of Recreation and Tourism in Natural Areas: A Case Study in New South Wales, Australia », *Tourism Economics*, vol. 3, n°3, pp. 227-239.

Carlsen J. et Wood D. (2004), Assessment of the economic value of recreation and tourism in Western Australia's national parks, marine parks and forests, Gold Coast, Qld, CRC for Sustainable Tourism, 29 p.

Caron N. (2005), « La correction de la non-réponse par repondération et par imputation », Insee document de travail, n°M0502, Paris, INSEE.

Cochran W.G. (1977), *Sampling Techniques*, Third Edition, New York, John Wiley, 442 p.

Cordell H.K., Bergstrom J.C. et Watson A.E. (1992), « Economic Growth and Interdependence Effects of State Park Visitation in Local and State Economies », *Journal of Leisure Research*, vol. 24, n°3, pp. 253-268.

Croes R.R. et Severt D.E. (2007), « Research Report: Evaluating Short-Term Tourism Economic Effects in Confined Economies – Conceptual and Empirical Considerations », *Tourism Economics*, vol. 13, n°2, pp. 289-307.

Crompton J.L. (2006), « Economic Impact Studies: Instruments for Political Shenanigans? », *Journal of Travel Research*, vol. 45, n°1, pp. 67-82.

Crompton J.L. (2010), « Measuring the Economic Impact of Park and Recreation Services », Ashburn, National Recreation and Park Association.

Crompton J.L., Lee S. et Shuster T.J. (2001), « A Guide for Undertaking Economic Impact Studies: The Springfest Example », *Journal of Travel Research*, vol. 40, n°1, pp. 79-87.

Davezies L. (2000), « Revenu et territoires », in *Aménagement du territoire: rapports*, Paris, Documentation Française (Les rapports du Conseil d'Analyse Economique), pp. 173-192.

Davezies L. (2004), « Formes de développement des territoires et pauvreté », pp. 21.

Davezies L. (2009), « L'économie locale « résidentielle » », *Géographie, économie, société*, vol. 11, n°1, pp. 47-53.

Davezies L. (2010), « La crise et nos territoires : premiers impacts », AdCF.

Deroyon T. (2018), « La correction de la non-réponse par repondération », Fiches méthodologiques, INSEE, Département des méthodes statistiques.

Deroyon T. et Favre-Martinoz C. (2018), « La correction de la non-réponse par imputation », Fiches méthodologiques, INSEE, Département des méthodes statistiques.

Deville J.-C. et Lavallée P. (2006), « Sondage indirect : Les fondements de la méthode généralisée du partage des poids », vol. 32, n°2, pp. 185-196.

Deville J.-C. et Maumy M. (2005), « Extensions de la méthode d'échantillonnage indirect et son application à l'enquête dans le tourisme: M.O.R.G.O.A.T », pp. 15.

Deville J.-C. et Maumy M. (2006), « Extensions de la méthode d'échantillonnage indirect et son application aux enquêtes dans le tourisme. Techniques d'enquête », vol. 32, n°2, pp. 197-206.

Driml S. (2010), « The economic value of tourism to national parks and protected areas in Australia », technical report.

Driml S. et McLennan C.-L. (2010), « Handbook on Measuring the Economic Value of Tourism to National Parks », Report, Sustainable Tourism Cooperative Research Centre.

Dwyer L., Forsyth P. et Spurr R. (2004), « Evaluating tourism's economic effects: new and old approaches », *Tourism Management*, vol. 25, n°3, pp. 307-317.

English D.B.K., Kocis S.M., Zarnoch S.J. et Arnold J.R. (2002), « Forest Service National Visitor Use Monitoring Process: Research Method Documentation », *Gen. Tech. Rep. SRS-57. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 14 p.*, vol. 057.

English D.B.K., Kriesel W., Leeworthy V.R. et Wiley P.C. (1996), « Economic contribution of recreating visitors to the Florida Keys/Key West », *Silver Spring, MD: National Oceanic and Atmospheric Administration, Strategic Environmental Assessments Division. 22 p.*

Eurostat (2000), *Methodological Manual on the Design and Implementation of Surveys on Inbound Tourism*, 2000^e édition, Office for Official Publications of the European Communities, 194 p.

Frechtling D.C. (1994), « Assessing the economic impacts of travel and tourism – Introduction to travel economic impact estimation. Second edition », *in Travel, Tourism and Hospitality Research: A Handbook for Managers and Researchers*, J.R. Brent Ritchie and Charles R. Goeldner, New York, John Wiley and Sons Inc., pp. 16.

Frechtling D.C. et Horváth E. (1999), « Estimating the Multiplier Effects of Tourism Expenditures on a Local Economy through a Regional Input-Output Model », *Journal of Travel Research*, vol. 37, n°4, pp. 324-332.

- Gazel R.C. et Schwer R.K. (1997)**, « Beyond Rock and Roll: The Economic Impact of the Grateful Dead on a Local Economy », *Journal of Cultural Economics*, vol. 21, n°1, pp. 41-55.
- Harnik P. et Crompton J.L. (2014)**, « Measuring the total economic value of a park system to a community », *Managing Leisure*, vol. 19, n°3, pp. 188-211.
- Henry E. et Deane B. (1997)**, « The contribution of tourism to the economy of Ireland in 1990 and 1995 », *Tourism Management*, vol. 18, n°8, pp. 535-553.
- Hjerpe E.E. et Kim Y.-S. (2007)**, « Regional economic impacts of Grand Canyon river runners », *Journal of Environmental Management*, vol. 85, n°1, pp. 137-149.
- Hughes M., Jones T., Deery M., Wood D., Fredline L., Whitely Z. et Lockwood M. (2009)**, « Estimating the Economic, Social, and Environmental Value of Tourism to Protected Areas », Technical Reports.
- Huhtala M. (2007)**, « Assessment of the local economic impacts of national park tourism: the case of Pallas-Ounastunturi National Park », *For. Snow Landsc. Res.*, pp. 16.
- Huhtala M., Kajala L. et Vatanen E. (2010)**, « Local economic impacts of national park visitors' spending in Finland: The development process of an estimation method », pp. 20.
- Jeong J.Y. et Crompton J.L. (2015)**, « Measuring the economic impact of a state park system », *Managing Sport and Leisure*, vol. 20, n°4, pp. 238-257.
- Jeong J.Y., Crompton J.L. et Dudensing R.M. (2016)**, « The Potential Influence of Researchers' "Hidden" Procedure Decisions on Estimates of Visitor Spending and Economic Impact », *Journal of Travel Research*, vol. 55, n°7, pp. 874-888.
- Job H. (2008)**, « Estimating the regional economic impact of tourism to national parks - two case studies from Germany », *GAIA - Ecological Perspectives for Science and Society*, vol. 17, n°1, pp. 134-142.
- Johnson R.L. et Moore E. (1993)**, « Tourism impact estimation », *Annals of Tourism Research*, vol. 20, n°2, pp. 279-288.
- Jones T. et Wood D. (2008)**, « Research Note: Attribution, Substitution and Confusion in Wildlife and Natural Area Tourism Expenditure Research », *Tourism Economics*, vol. 14, n°2, pp. 431-434.
- Kalton G. (1991)**, « L'Échantillonnage des flux de populations humaines mobiles », vol. 17, n°2, pp. 197-210.
- Kalton G. (2001)**, « Practical Methods for Sampling Rare and Mobile Populations », pp. 6.
- Kemper N.P., Popp J.S. et Miller W.P. (2008)**, « Regional Growth and Beaver Lake: A Study of Recreation Visitors », *Tourism Economics*, vol. 14, n°2, pp. 409-426.
- Kirillova K., Bordelon B.M. et Pearlman D. (2014)**, « A Comparison of Expenditures Between New Orleans Volunteer and Leisure Tourists: implications for Sustainability », *Tourism analysis*, vol. 19, n°2, pp. 173-184.

- Kumar J. et Hussain K. (2014)**, « Evaluating Tourism's Economic Effects: Comparison of Different Approaches », *Procedia - Social and Behavioral Sciences*, vol. 144, pp. 360-365.
- Lavallée P. (1995)**, « Pondération transversale des enquêtes longitudinales menées auprès des individus et des ménages à l'aide de la méthode du partage des poids », *Statistique Canada*, vol. 21, n°1, pp. 27-35.
- Lemel Y. (1976)**, « Une généralisation de la méthode du quotient pour le redressement des enquêtes par sondage », *Annales de l'insée*, n°22/23, pp. 273-282.
- Leroy C. (2018)**, « Méthodologie et documentation pour la base des TCM empilés », pp. 51.
- Lindberg K. et Denstadli J.M. (2004)**, Impact of national park visitation on rural economies and government revenue in Queensland: examples of Girraween, Eungella, Daintree and Carnarvon, Gold Coast, Qld., CRC for Sustainable Tourism.
- Long P.T. et Perdue R.R. (1990)**, « The Economic Impact Of Rural Festivals And Special Events: Assessing The Spatial Distribution Of Expenditures », *Journal of Travel Research*, vol. 28, n°4, pp. 10-14.
- Loomis J. (2007)**, « Correcting for On-Site Visitor Sampling Bias When Estimating the Regional Economic Effects of Tourism », *Tourism Economics*, vol. 13, n°1, pp. 41-47.
- Loomis J. et Caughlan L. (2006)**, « The Importance of Adjusting for Trip Purpose in Regional Economic Analyses of Tourist Destinations », *Tourism Economics*, vol. 12, n°1, pp. 33-43.
- Loonis V. (2009)**, « L'échantillonnage de la théorie à la pratique », pp. 8.
- Loveridge S. (2004)**, « A Typology and Assessment of Multi-sector Regional Economic Impact Models », *Regional Studies*, vol. 38, n°3, pp. 305-317.
- Lue C.-C., Crompton J.L. et Fesenmaier D.R. (1993)**, « Conceptualization of multi-destination pleasure trips », *Annals of Tourism Research*, vol. 20, n°2, pp. 289-301.
- Manente M. et Minghetti V. (1998)**, « Surveying visitors and tourist itineraries. The case of the archaeological area of Pompeii and Herculaneum », pp. 30.
- Marpsat M. (2008)**, « The INED research on homelessness (1993-2008) », *Documents de travail*, INED.
- Marpsat M. et Razafindratsima N. (2010)**, « Les méthodes d'enquêtes auprès des populations difficiles à joindre: Introduction au numéro spécial », *Methodological Innovations Online*, vol. 5, n°2, pp. 3.2-16.
- Martínez-Espiñeira R. et Amoako-Tuffour J. (2009)**, « Multi-Destination and Multi-Purpose Trip Effects in the Analysis of the Demand for Trips to a Remote Recreational Site », *Environmental Management*, vol. 43, n°6, pp. 1146-1161.
- Maurence E. (2012)**, « La mesure de l'impact économique d'un événement touristique », Rapport d'étude, EMC/Ministère de l'Économie, des Finances et de l'Industrie, DGCIS.

Mayer M., Müller M., Woltering M., Arnegger J. et Job H. (2010), « The economic impact of tourism in six German national parks », *Landscape and Urban Planning*, vol. 97, n°2, pp. 73-82.

McDonald G.T. et Wilks L.C. (1986), « The Regional Economic Impact of Tourism and Recreation in National Parks », *Environment and Planning B: Planning and Design*, vol. 13, n°3, pp. 349-366.

McGrath J.M., Primm D. et Lafe W. (2017), « Heritage tourism's economic contribution: A Pennsylvania case study », *Tourism Economics*, vol. 23, n°5, pp. 1131-1137.

McHone W.W. et Rungeling B. (2000), « Practical Issues in Measuring the Impact of a Cultural Tourist Event in a Major Tourist Destination », *Journal of Travel Research*, vol. 38, n°3, pp. 300-303.

Miller R.E. et Blair P.D. (2009), *Input–Output Analysis: Foundations and Extensions*, Second Edition, Second Edition.

Mules T. (2005), « Economic Impacts of National Park Tourism on Gateway Communities: The Case of Kosciuszko National Park », *Tourism Economics*, vol. 11, n°2, pp. 247-259.

Mules T. et Faulkner B. (1996), « An economic perspective on special events », *Tourism Economics*, vol. 2, n°2, pp. 107-117.

Mules T., Faulks P., Stoeckl N. et Cegielski M. (2005), « The Economic Values of Tourism in the Australian Alps », Report, Gold Coast, QLD, Cooperative Research Centre for Sustainable Tourism.

Nicolas Y. (2007), « Les premiers principes de l'analyse d'impact économique local d'une activité culturelle », *Culture méthodes*, n°1, pp. 1-8.

OCDE (2002), Manuel d'évaluation de la biodiversité.

Oosterhaven J. et Knijff E.C. van der (1987), « On the Economic Impacts of Recreation and Tourism: The Input-Output Approach », *Built Environment (1978-)*, vol. 13, n°2, pp. 96-108.

Orens A. et Seidl A. (2009), « Working Lands and Winter Tourists in the Rocky Mountain West: A Travel Cost, Contingent Behaviour and Input–Output Analysis », *Tourism Economics*.

Pearce D., Atkinson G. et Mourato S. (2007), *Analyse coûts-bénéfices et environnement. Développements récents*, OECD Publishing, 356 p.

Pecqueur B. et Talandier M. (2011), « Les territoires à base économique résidentielle et touristique », *TERRITOIRES 2040*.

Prideaux B. et Falco-Mammone F. (2007), *Economic values of tourism in the Wet Tropics World Heritage Area*, Cairns, Qld, Cooperative Research Centre for Tropical Rainforest Ecology and Management, 64 p.

Réseau des Grands Sites de France (2008), « Evaluer les retombées économiques des grands sites. », *Le fil des Grands Sites*, n°3.

Roy G. et Vanheuverzwyn A. (2001), « Redressement par la macro CALMAR : applications et pistes d'amélioration ».

SPOT Observatoire Tourisme / INSEE Auvergne (2011), « L'emploi salarié touristique en Auvergne »,.

Statistique Canada (2003), *Méthodes et pratiques d'enquête*, Canada, Ottawa, Statistique Canada, 396 p.

Stoeckl N., Smith A. et Newsome D. (2005), « Regional Economic Dependence on Iconic Wildlife Tourism: Case studies of Monkey Mia and Hervey Bay », pp. 13.

Stynes D.J. (1997a), « Economic Impacts of Tourism », pp. 19.

Stynes D.J. (1997b), « Economic impacts of tourism: A handbook for tourism professionals », Illinois Bureau of Tourism - Illinois Department of Commerce and Community Affairs.

Stynes D.J. (1999), « Guidelines for Measuring Visitor Spending », pp. 9.

Stynes D.J. (2002), Economic Impacts of Great Smoky Mt. National Park Visitors on the Local Region, 1997-2000.

Stynes D.J. (2007), « Impacts of Visitor Spending on the Local Economy: Yosemite National Park, 2005 ».

Stynes D.J. et Sun Y.-Y. (2003), « Economic Impacts of National Park Visitor Spending on Gateway Communities : Systemwide Estimates for 2001 ».

Stynes D.J. et White E.M. (2006), « Reflections on Measuring Recreation and Travel Spending », *Journal of Travel Research*, vol. 45, n°1, pp. 8-16.

Stynes D.D.J. (2005), « Economic Significance of Recreational Uses of National Parks and Other Public Lands », vol. 5, n°1, pp. 36.

Sun Y.-Y. et Stynes D.J. (2006), « A note on estimating visitor spending on a per-day/night basis », *Tourism Management*, vol. 27, n°4, pp. 721-725.

Talandier M. (2008a), « Richesse et développement des territoires », *Pour*, vol. n° 199, n°4, pp. 53-59.

Talandier M. (2008b), « Le tourisme: moteur de la revitalisation des espaces ruraux français? », pp. 132-136.

Talandier M. (2008c), « Une autre géographie du développement rural : une approche par les revenus », *Géocarrefour*, vol. 83, n°Vol. 83/4, pp. 259-267.

Talandier M. (2012), « L'émergence des systèmes productivo-résidentiels », pp. 25.

Tillé Y. (1999), « Utilisation d'informations auxiliaires dans les enquêtes par sondage », *Questiio: Quaderns d'Estadística, Sistemes, Informàtica i Investigació Operativa*, ISSN 0210-8054, Vol. 23, N°. 3, 1999, pags. 491-505, vol. 23.

Tisdell C.A. et Wilson C. (2002), « Economic, Educational and Conservation Benefits of Sea Turtle Based Ecotourism: A Study focused on Mon Repos ».

Tremblay P. (2007), Economic contribution of Kakadu National Park to tourism in the Northern Territory, Gold Coast, Qld. : Sustainable Tourism CRC.

Tremblay P. et Carson D.B. (2007), Tourism and the economic valuation of parks and protected areas: Watarrka National Park, Northern Territory, Gold Coast, Qld., CRC for Sustainable Tourism Pty Ltd.

Tyrrell T.J. et Johnston R.J. (2001), « A Framework for Assessing Direct Economic Impacts of Tourist Events: Distinguishing Origins, Destinations, and Causes of Expenditures », *Journal of Travel Research*, vol. 40, n°1, pp. 94-100.

Tyrrell T.J. et Johnston R.J. (2016), « The Economic Impacts of Tourism: A Special Issue », *Journal of Travel Research*.

Vanderschelden M. (2005), « Homogamie et choix du conjoint - Traitement de la non-réponse. Imputation de variables qualitatives corrélées », Document de travail de la Direction des Statistiques Démographiques et Sociales de l'Insee, n°F0505, Insee.

Warnick R.B., Bojanic D.A. et Xu F. (2012), « Economic Impact and Significance: Additional Insights for Measurement in Special Events », pp. 8.

Warnick R.B., Bojanic D.C. et Cartier E. (2017), « A Comparison of Economic Impact Measurement Techniques for a Tourism Special Event », *Journal of Travel Research*, vol. 56, n°5, pp. 578-592.

Warnick R.B., Bojanic D.C., Mathur A. et Ninan D. (2011), « Segmenting Event Attendees Based on Travel Distance, Frequency of Attendance, and Involvement Measures: A Cluster Segmentation Technique », *Event Management*, vol. 15, n°1, pp. 77-90.

Warnick R.B., Bojanic D.C. et Xu F. (2015), « Using a Trade Market Analysis Technique to Refine Measurements for Economic Impact Analysis of Special Events », *Journal of Travel Research*.

Warnick R., Bojanic D. et Cartier E. (2016), « Economic Impact Research: Review of Measurement Techniques in a Special Event », pp. 14.

Watson P., Thilmany D., Wilson J. et Winter S. (2007), « Determining Economic Contributions and Impacts: What is the difference and why do we care? », *Journal of Regional Analysis and Policy*, vol. 37, pp. 140-146.

Wood D., Glasson J., Carlsen J., Hopkins D., Terry P., Lacy D., Tourism S., Executive C.C., Leo P., Sustainable J. et Crc T. (2006), *Economic Evaluation of Tourism for Natural Areas: Development of a Toolkit Approach*.

VII. Bibliographie complémentaire

Aznar, O. (2004). Services environnementaux et espaces ruraux. PhD Thèse de doctorat, Université de Bourgogne.

Binet, T., A. B. de Battisti and P. Failler (2011). "Évaluation économique des écosystèmes marins et côtiers." Jacquet, P., Kumar, R. and L. Tubiana (eds).

Bonniot, A., M. Bouchaud, M. Guillot, J.-B. Marsat, H. Blasquiet-Revol and A. Gaunot (2010). De l'importance de l'implication des acteurs privés dans la pérennisation d'une destination touristique. Illustration à partir du cas de La Chaise Dieu et de son festival de musique sacrée. XLVII^e Colloque de l'Association de Science Régionale de Langue Française (ASRDLF), Aoste, ASRDLF.

CGDD (2010). Donner une valeur à l'environnement : la monétarisation, un exercice délicat mais nécessaire.

CREDOC (2008). "La valeur économique et sociale des espaces naturels protégés." Cahier de Recherche du Credoc(247): 103.

CREDOC (2008). Les retombées économiques et les aménités des espaces naturels protégés. Paris, CREDOC Département "évaluation des Politiques Publiques": 131p.

CRT_RA (2016). C a r n e t d e l ' o b s e r v a t i o n l o c a l e d u t o u r i s m e. Lyon.

Duval, M., B. Smith, S. Høerlé, L. Bovet, N. Khumalo and L. Bhengu (2019). "Towards a holistic approach to heritage values: a multidisciplinary and cosmopolitan approach." International Journal of Heritage Studies: 1-23.

Eiglier, P. and E. Langeard (1987). Servuction - le marketing des services. Paris, Ediscience international.

EUROSTAT (2000). Methodological Manual on the Design and Implementation of Surveys on Inbound Tourism, Office for Official Publications of the European Communities.

Gadrey, J. (1994). L'économie des services. Paris, La Découverte.

Marsat, J.-B. and A. Bonniot (2010). Penser un "tourisme territorial intégré" et sa gouvernance – Modèles et pratiques. L'écotourisme visité par les acteurs territoriaux. C. Gagnon and S. Gagnon. Montréal, PUQ Presses Universitaires du Québec: 93-120.

Marsat, J.-B., A. Bonniot, M. Bouchaud and C. Monin (2013). Territorial Integrated Tourism : dispersion, complementarity, social embeddedness and governance - conceptual framework and evidence in Auvergne (France). The Sustainability of Rural Systems: Local and Global Challenges and Opportunities. M. Cawley, A. M. Bicalho and L. Laurens. Galway, Whitaker Institute, NUI Galway and CSRS of the International Geographical Union: pp. 245-276.

Marsat, J.-B., S. Brault, B. Dazet and C. Lacroix (2009). Le management territorialisé du tourisme - Des enseignements à partir du programme Porta Natura. Destinations et territoires. J.-P. Lemasson and P. Violier. Montréal--Québec, UQAM - ESG- TEOROS. 1: 168-177.

Meyronin, B. and J.-P. Valla (2004). Vers la notion de « servuctions urbaines », ou les apports du marketing des services au marketing territorial. Congrès Paris-Venise -4ème Congrès Tendances du Marketing, Venise, AIMS.

Min._Economie and E. Maurence (2012). La mesure de l'impact économique d'un événement touristique. Rapport d'étude, Paris, ministère de l'Économie, des Finances et de l'Industrie, DGCIS, Sous-direction de la Prospective, des Études économiques et de l'Évaluation.

Ministère_environnement (2017). Comment évaluer les valeurs économique, sociale, environnementale et territoriale ? Une méthode opérationnelle pour les acteurs des territoires d'exception. Paris, Ministère de l'environnement de l'énergie et de la mer: 184 p.

Petiot, R. (2004). "Réflexions sur la question de la création de la valeur touristique." Téoros. Revue de recherche en tourisme 23(23-3): 50-54.

Phetvaroon, K. (2006). Application of the theory of planned behavior to select a destination after a crisis: A case study of phuket, Thailand. PhD, University of Tennessee.

UNESCO (2021). Visitors count! Guidance for protected areason the economic analysis of visitation, UNESCO.

VIII. Annexes

A. Annexe : Caractéristique des zones d'impact

| Données 2017 | Allier | Cantal | Haute-Loire | Puy-de-Dôme | France métropolitaine | Allier/F | Cantal/F | Haute-Loire/F | Puy-de-Dôme/F |
|---|---------|---------|-------------|-------------|-----------------------|----------|----------|---------------|---------------|
| Superficie (km2) | 7 340 | 5 726 | 4 977 | 7 970 | 543 940 | 0,01 | 0,01 | 0,01 | 0,01 |
| Population | 337 988 | 145 143 | 227 283 | 653 742 | 64 639 133 | 0,01 | 0,00 | 0,00 | 0,01 |
| Densité (hab / km2) | 46,00 | 25,30 | 45,70 | 82,00 | 118,80 | 0,39 | 0,21 | 0,38 | 0,69 |
| Variation annuelle Totale | -0,30 | -0,30 | 0,10 | 0,50 | 0,40 | -0,75 | -0,75 | 0,25 | 1,25 |
| Variation Naturelle de la population (2012 et 2017, en %) | -0,40 | -0,60 | -0,20 | 0,10 | 0,30 | -1,33 | -2,00 | -0,67 | 0,33 |
| % de la population ≥ 60 ans | 34,10 | 35,20 | 30,70 | 27,80 | 25,70 | 1,33 | 1,37 | 1,19 | 1,08 |
| Actifs ayant un emploi | 62,70 | 69,00 | 67,10 | 64,4 | 64,20 | 0,98 | 1,07 | 1,05 | 1,00 |

| | | | | | | | | | | |
|--------------------------------|---|--------|--------|--------|--------|--------|------|------|------|------|
| | Chômeurs | 10,20 | 6,70 | 7,90 | 8,90 | 9,90 | 1,03 | 0,68 | 0,80 | 0,90 |
| % dans la population 15-64 ans | Élèves, étudiants et stagiaires non rémunérés | 8,00 | 7,20 | 8,50 | 11,50 | 10,60 | 0,75 | 0,68 | 0,80 | 1,08 |
| | Retraités et pré-retraités | 9,50 | 9,80 | 9,10 | 7,60 | 6,80 | 1,40 | 1,44 | 1,34 | 1,12 |
| | Autres inactifs | 9,60 | 7,40 | 7,30 | 7,50 | 8,50 | 1,13 | 0,87 | 0,86 | 0,88 |
| | Actifs ayant un emploi / population totale en % | 36,40 | 40,27 | 39,98 | 40,26 | 40,05 | 0,91 | 1,01 | 1,00 | 1,01 |
| Structure de l'emploi | Agriculture | 5,30% | 12,10% | 6,30% | 3,00% | 2,60% | 0,02 | 4,65 | 2,42 | 1,15 |
| | Industrie | 15,20% | 10,50% | 19,90% | 15,70% | 12,20% | 0,01 | 0,86 | 1,63 | 1,29 |
| | Construction | 6,40% | 8,00% | 7,80% | 6,30% | 6,40% | 0,01 | 1,25 | 1,22 | 0,98 |
| | Services marchands | 36,90% | 33,80% | 32,00% | 41,60% | 46,80% | 0,01 | 0,72 | 0,68 | 0,89 |

| | | | | | | | | | |
|--|--------|--------|-------|--------|--------|------|------|------|------|
| Services non- marchands et santé | 36,20% | 35,50% | 34,10 | 33,50% | 32,00% | 0,01 | 1,11 | 1,07 | 1,05 |
| <hr/> | | | | | | | | | |
| Salaire net horaire moyen (€) | 12,45 | 11,72 | 12,22 | 13,92 | 14,89 | 0,84 | 0,79 | 0,82 | 0,93 |
| <hr/> | | | | | | | | | |
| Taux de résidences secondaires en % | 7,30 | 20,40 | 16,30 | 10,20 | 9,80 | 0,74 | 2,08 | 1,66 | 1,04 |

B. Annexe : Les traitements des données

Les éléments en rouge dans les fichiers fusionnés sont des modifications apportées ou des choses à modifier.

- Les colonnes : Test Restauration ; Test Alimentation ; Test Souvenir et Produits locaux ; Test visites ou activités sportives, culturelle, événements et achats d'équipements ; Test souvenir ou d'équipement (VERSION1) ; Test produits locaux et l'alimentation (VERSION1) ; Test Activités sportives, culturelle, événements (VERSION1) ; Test budget total sont des colonnes de comparaison des dépenses journalières et du séjour des touristes.
 - 0 = le budget est correct
 - 1 = le budget journalier est supérieur au budget du séjour
 - 2 = soit le budget journalier soit le budget du séjour sont inexistants.

Fichiers bruts

A chaque retour d'enquête ou au lendemain d'une période d'enquête, Boris (quelquefois Alexia) se charge de télécharger les Raspberry (3 au total) utilisés. Les données recueillies sont importées de LimeSurvey (via les Raspberry) vers Excel. Ainsi, à chaque Raspberry utilisé pendant une journée d'enquête correspond un fichier Excel.

Fichier Traité

Fusion des fichiers et données de traçage

Une fois les données importées dans des fichiers Excel, on procède à la fusion de ces derniers. Ainsi, pour faciliter cette fusion des différents fichiers issus des enquêtes menées pendant différentes périodes ou par différents enquêteurs, nous avons procédé à la création de 5 cinq colonnes (N°indiv, Date_Enquete, Semaine, Mois, Questionnaire LimeSurvey) dans le fichier Excel.

- **N°indiv** : un numéro est attribué à chaque observation (chaque personne ou ménage enquêté). Ce numéro permet d'identifier chaque observation. En effet, les « ID de la réponse » issus d'un Raspberry peuvent correspondre à ceux d'un autre Raspberry pour une même journée d'enquête. Aussi, l'évolution du questionnaire d'enquête au cours de la période d'enquête et le fait de télécharger au fur et à mesure les données recueillies, ont conduit à mettre lors de certaines sorties d'enquête un questionnaire modifié ou non sur les Raspberry.
- **Date_Enquete** : une date est attribuée à chaque observation correspondant à la date d'enquête.
- **Semaine** : un numéro de semaine selon la date d'enquête (selon le calendrier grégorien.) est attribué à chaque observation.
- **Mois** : Correspond au mois d'enquête.
- **Questionnaire LimeSurvey** : la version du questionnaire utilisé.

Les fichiers Excel ainsi fusionnés sont mis dans le dossier nommé « **Fichiers traités** », lui-même situé dans le dossier nommé « Résultats export LimeSurvey » le tout est placé dans le « Datafiles ». Tous les fichiers Excel fusionnés sont nommés comme suit : « période d'enquête concernée + Traité ».

Quelques corrections effectuées


Au fur et à mesure que les fichiers Excel sont fusionnés, et ensuite, après la fin de l'enquête, nous procédons à certaines corrections des données recueillies. Ces corrections correspondent généralement :

- Aux erreurs de frappe
- A des permutations involontaires, par exemple entre deux rubriques consécutives : par exemple « quelles sont vos dépenses du jour XXX, », et « quelles sont vos dépenses du séjour en XXX » : dans certains cas évidents d'inversion, les données sont rectifiées
- A d'autres cas particuliers, à partir d'informations déclarées en commentaire, ou notées par l'enquêteur

D'autres traitements sont effectués ensuite

- Reconstitution de certaines dépenses.
 - Par exemple, certains visiteurs ne se souviennent pas exactement du montant payé lors de la visiter d'un ou plusieurs sites payants. Nous procédons donc à la reconstitution de ces montants en se basant sur les prix d'entrée de ces sites et de l'effectif du groupe concerné. Dans une majorité de cas, les enquêtés nous ont énuméré les sites visités.
 - De même lorsque les valeurs pour l'alimentation sont absentes, elles ont été estimées à partir de valeurs unitaires moyennes (ou médianes). Cela n'est pas fait lorsque l'enquêté a déclaré explicitement une valeur nulle (par exemple dans le cas d'alimentation apportée par un touriste depuis son domicile).
 - De même une dépense de transport pour le jour de l'enquête a été estimée à partir des distances entre le lieu d'hébergement (permanent pour les excursionnistes, temporaire pour les touristes) et un point « central » de la Chaîne des Puys
- Décomposition de certaines dépenses : cas des « pack », pensions, forfaits : Lorsque c'est possible ces groupes de dépenses ont été décomposés
 - Par exemple le forfait « hébergement + entrée à Vulcania » a pu être décomposé connaissant les tarifs de Vulcania (une approximation doit être opérée car les tarifs sont sans doute affectés d'une réduction dans ce cadre)

Couleurs utilisées

- Au cours de l'enquête, certaines questions sont ajoutées, modifiées ou supprimées du questionnaire. Ainsi, la couleur « grise »  visible dans certaines colonnes des fichiers correspond aux périodes où la question concernée n'est pas posée.
- En ce qui concerne les colonnes « dépenses »
 - Les colonnes « Calcul » **en couleur verte** : les valeurs dans ces colonnes correspondent au centre des intervalles des dépenses correspondantes.

- Les colonnes « Différence » **en couleur verte** : les valeurs dans ces colonnes correspondent à la différence entre **les budgets du séjour et les budgets du jour**. Cette différence est effectuée à partir des données des colonnes « calcul ». Une valeur négative dans cette colonne correspond :
 - + ✚ Soit à un excursionniste qui n'a pas de budget séjour (donc c'est normal)
 - + ✚ Soit à un touriste pour qui les dépenses du séjour n'ont pas été déclarées ou la dépense du séjour est inférieure à la dépense du jour. Dans ce cas une correction doit être effectuée. Ainsi, on peut considérer que la dépense du jour correspond à la dépense du séjour, ou reconstruire cette dépense à partir des données disponibles ou permuter les deux dépenses.
- Les colonnes « Test » **en couleur bleu** : dans ces colonnes :
 - + ✚ 0 signifie que le budget du séjour et du jour sont correctes
 - + ✚ 1 signifie que le budget du séjour **est inférieur** au budget du jour (donc une correction s'impose)
 - + ✚ 2 signifie que les budgets du séjour et du jour doivent être vérifiés.
 - Les colonnes de calcul en couleurs rose ou rouge :
 - En rose sont les calculs intermédiaires
 - En rouge les valeurs finales

● Exemple au 22 Juillet 2020 de la feuille de calcul « Hbgt pension pack »

- Les colonnes de 1 à 37 sont identiques à celles qui sont dans les autres feuilles de calcul
 - Elles sont principalement issues des réponses au questionnaire sur des questions importantes pour le traitement de tous les types de dépense (comme : êtes-vous touriste ou excursionniste ? », ou sur des questions particulièrement importantes pour le thème de la feuille en cours : par exemple « quel est votre type d'hébergement ? » etc.
 - Parmi les premières colonnes on trouve des indicateurs créés pour le traitement (code selon que les données sont complètes ou non, commentaires, etc.)
- Puis les colonnes 38 à 68 sont spécifiques au thème (hébergement) ou nécessaires à son traitement (dépenses de restaurant pour décomposer les cas de pension) enfin on a gardé les réponses des enquêtés qui répondent en termes de budget global pour être capable de vérifier la cohérence de ces cas (ils peuvent avoir déclaré aussi quelques dépenses de détail)
- Illustration

| Intitulé | Contenu |
|---|---|
| Pouvez-vous indiquer votre budget global concernant l'hébergement ? | Réponse brute de l'interviewer Soit par plages de valeurs Soit en valeur précise (que nous avons reportée à partir de la dernière colonne ci-dessous) |
| Calcul hébergement valeurs | Conversion de la plage de valeurs (ici par valeur médiane) |
| Calcul hébergement valeurs Scén 1 : "plus de " = tranche suppl | Calcul alternatif proposant une autre valeur pour les cas « plus de 900 euros |
| Calcul hébergement valeurs Scén 1 : "plus de " = 1000 valeurs seules | Identique à la précédente en supprimant les valeur alphanumériques (messages d'erreur, etc.) |
| Calcul hébergement Scén 1 : "plus" de = 1000 valeurs par p et par nuit | Ratio simple : dépense par personne et par nuit |
| Calc Hbgt val moyenne par p et par nuit Scén 1 par type d'hbgt | Ratio 2 : ici on note le résultat des moyennes du ratio par personne et par nuit, moyennes calculées après tri selon le type d'hébergement |
| Calc Hbgt simul ratio X effectif X durée | Simulation par application du ratio précédent à chaque séjour (nombre de personnes et durée) au sein d'un même type d'hébergement |
| Calc Hbgt repl valeurs "0" ou "plus de" estim p ratio / sc1 | Imputation pour les cas de valeurs absentes ou supérieures au plafond |

| | |
|---|--|
| par type d'hbgt av les 67 cas "budg global | NB : ce calcul n'est pas définitif (pas encore de traitement des pensions ou packs De plus ici on garde provisoirement ces calculs même dans les cas où l'enquêté a répondu en termes de budget global (car cela pourrait permettre d'étendre l'échantillon dans une analyse spécifique des dépenses d'hébergement) |
| Calc Hbgt rempl valeurs "0" ou "plus de" estim p ratio / sc1 par type d'hbgt sans cas Budg global | Idem sans les cas « budget global » (case vide) |
| Calc Hbgt Pension pack déduit restau et actis si pension ou pack | Dernier traitement : Les cas de pension ou de pack sont traités au mieux par estimation de chacune des deux composantes : hbgt d'une part, et d'autre part soit restauration (cas des pensions) soit activité payante (cas des packs) |
| Calc Hbgt Pension pack déduit restau et actis si pension ou pack valeurs num | Idem sans les valeurs alphanumériques (messages d'erreur) |
| Calc Hbgt déduit restau et actis si pension ou pack sans cas Budg global | Idem sans les cas « budget global » |
| Pouvez-vous indiquer votre budget global concernant l'hébergement ? [Autre] | Réponse brute de l'interviewé en valeur précise éventuellement avec commentaires |

C. Annexe : Structure des cellules du fichier exporté LimeSurvey vers Excel

N/A : elle est remplie automatiquement pour **les non-réponses** (soit par oubli, soit parce que la question est sans objet) selon la règle suivante :

- Question à réponses prédéfinies intégrées dans LimeSurvey telles que : Oui/Non, Homme/Femme.
- Question à choix multiple avec plusieurs modalités de réponses (participant peut sélectionner plusieurs réponses pour la même question) éclatée en plusieurs question. A noter que dans le cas de ce type de question, si une réponse (oui ou non) est donnée pour l'une des modalités, les autres modalités sont automatiquement mises à « non ».

Des vides : correspondent aux **non-réponses** (soit par oubli, soit parce que la question est sans objet) aux « questions non posées » éditée (c'est-à-dire non prédéfinie par LimeSurvey) par le concepteur du questionnaire

- **Question ouverte** : question à réponses non suggérées par le concepteur du questionnaire. L'enquêté répond librement, avec ses propres mots, dans une zone de texte prédéfinie.
- **Question à choix unique avec plusieurs modalités de réponses**. C'est le cas des réponses aux questions liées aux dépenses.

Dans les deux cas ces absences de réponses proviennent de situations différentes

- Question sans objet (cas des dépenses de séjour pour un excursionniste)
- Question non posée parce que la version du questionnaire en vigueur ne la prévoyait pas encore
- Question non posée par oubli de la part de l'enquêteur
- Question sans réponse du fait que l'enquêté ne sait pas
- Question sans réponse du fait que l'enquêté ne veut pas répondre

Activités effectuées :

Les feuilles Excel « Activités effectuées » dans la base de données recensent tous les sites/lieux payants/gratuits visités les enquêtés ; ou les activités payantes/gratuites effectuées par les enquêtés. Ces fichiers sont issus des fichiers traités. Ils nous permettent de calculer les dépenses liées aux billetteries de chaque visiteur. Dans chaque colonne, nous avons :

- 1 si l'enquêté a visité le site ou a effectué l'activité mentionnée
- 0 sinon

A noter que cette information est étendue à tout le groupe pour lequel l'enquêté répond.

D. Annexe : Quelques corrections/imputations effectuées dans le fichier

- 1) Budget du jour est supérieur au budget du séjour
 - Si le Budget du séjour = 0, le budget du jour est repris au niveau du Budget du séjour. Ainsi le Budget du séjour = Budget du jour.
 - Si le Budget du séjour > 0, le Budget du séjour et le Budget du jour sont permutés. Ainsi le Budget du séjour > Budget du jour.
- 2) Budget du séjour = 0 alors que le Budget du jour = NVPR, NSP, N/A
 - Budget du jour = 0
- 3) Combien de fois allez-vous venir sur le site de la CP sur l'ensemble de votre séjour ?
 - Si la cellule est vide et durée du séjour = 1 alors imputé "1"
 - Si c'est = 0 => N/A
- 4) Quel est votre type d'hébergement pour votre séjour ?
 - Si la réponse est « Hébergement chez de la famille, des amis en résidence principale », normalement la réponse à la question « Pouvez-vous indiquer votre budget global concernant l'hébergement ? » doit être = 0 euros. Mais il y a certaines observations qui ont donné des montants supérieurs à 0. On a : 683 ; 835 ; 1033 ; 1070. Il est convenu de conserver ces chiffres car ils peuvent traduire certaines participations aux frais d'accueil.
- 5) L'observation 360 est changée en excursionniste
- 6) L'observation 1041 à supprimer car c'est un test
- 7) Diriez-vous que vous êtes venu cette fois-ci dans cette destination / dans votre résidence secondaire ?
 - Toutes les réponses des excursionnistes sont transformées en « **sans objet (SO)** »
- 8) Si vous êtes venu en partie pour la CP, est-ce :
 - SO pour les observations suivantes : 150 ; 167 ; 176 ; 224 ; 282 ; 151 ; 223 ; 229
- 9) La durée du séjour de l'observation 29 est corrigé =1 vu les dépenses
- 10) Si hébergement temporaire, quel est votre mode de tourisme ?
 - **Créé le mode « étape » pour les cas** : 85 ; 115 ; 136 ; 172 ; 186 ; 211 ; 682 ; 699 ; 1040 ; 1033

- 11) Observation 45 : seul le lieu d'hébergement (département) est changé
- 12) Les budgets indiqués ci-dessous sont-ils pour l'ensemble de votre famille / groupe / couple ou seulement pour vous ?
- Corrigé à partir de "Effectif du groupe concerné par le budget corrigé V 17/4"
- 13) Si hébergement temporaire, quel est votre mode de tourisme ?
- Observation 1033 « hébergé chez de la famille... » => Hébergé en un seul lieu

E. Annexe : Précisions mathématiques : moyennes, variance pondérée

Moyennes :

Concernant les données de chaque enquêtes qui sont proportionnelles aux effectifs du groupe enquêté, aux durées de séjour, etc. il existe deux méthodes pour calculer des moyennes :

- Méthode « somme / somme » : Par exemple somme des dépenses totales de tous les groupes / somme des effectifs de tous les groupes
- Méthode « moyenne de moyenne » : Par exemple moyenne des moyennes de dépense par personne calculées pour chaque groupe

Les différences de résultats obtenus avec les deux méthodes peuvent être très sensibles (jusqu'à 30% au moins) selon les distributions de données. En général il semble bien que la méthode somme/somme donne les résultats inférieurs et la moyenne des moyennes donne des résultats supérieurs. C'est que signale la littérature consultée, sans donner de preuve mathématique, et c'est ce que nous avons constaté par des tests sur quelques unes des données de l'enquête.

Sauf cas particulier explicité (pour comparer les résultats avec des données qui ont été calculées par la moyenne des moyennes, comme celles du SDT), la présente étude utilise la méthode « somme/somme ».

Les résultats globaux sont ainsi bien inférieurs à ce que donnerait l'application systématique de l'autre méthode (dans les phases d'imputation, et la phase finale de calcul des moyennes de dépense par catégorie de visiteur).

Variance pondérée

Les données de chaque enquête sont affectée d'un coefficient de redressement. De plus la plupart des variables sont affectées d'une pondération par les effectifs de chaque groupe enquêté, ou par des durées. Ainsi les calculs de variance et écart-type doivent prendre la forme de calculs « pondérés ». La formule appliquée est la suivante :

formule Ecart-type "regroupé" ou "pondéré"

$$\sigma^2 = \frac{\sum_{i=1}^N w_i (x_i - \bar{x}^*)^2}{\sum_{i=1}^N w_i}$$

x barre * est la moyenne pondérée