



HAL
open science

Zones difficiles et notoriété touristique. L'apport de l'analyse textuelle dans la construction d'indicateurs numériques décrivant des processus spatiaux. Le cas de la notoriété touristique des communes corses

Claude Napoleone, Guillaume Ollivier, Ghislain Geniaux, Jean Christophe Paoli, Johan Jouve

► **To cite this version:**

Claude Napoleone, Guillaume Ollivier, Ghislain Geniaux, Jean Christophe Paoli, Johan Jouve. Zones difficiles et notoriété touristique. L'apport de l'analyse textuelle dans la construction d'indicateurs numériques décrivant des processus spatiaux. Le cas de la notoriété touristique des communes corses. Colloque International Sociétés en transition et développement local en zones difficiles DELZOD, Apr 2009, Medenine, Tunisie. hal-02751444

HAL Id: hal-02751444

<https://hal.inrae.fr/hal-02751444>

Submitted on 3 Jun 2020

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.

Zones difficiles et notoriété touristique

L'apport de l'analyse textuelle dans la construction d'indicateurs numériques décrivant des processus spatiaux. Le cas de la notoriété touristique des communes Corses

Napoléone Claude¹, Ollivier Guillaume², Geniaux Ghislain¹, Jouve Johan³ et Paoli Jean-Christophe³

Le tourisme de masse est un fait marquant de nombreux pays méditerranéens. A titre d'exemple, si la France reçoit à peu près un touriste par habitant, la région PACA (France) reçoit 30 millions de visiteurs par an alors qu'elle est peuplée de 5 millions de personnes et la Corse plus de 2 millions de touristes pour moins de 300000 habitants. De telles fréquentations suscitent une organisation quasi industrielle qui, dès lors qu'elle se développe dans un environnement économique et social peu dynamique, des zones difficiles pour reprendre la terminologie du colloque, est à même de modifier en profondeur la physionomie ou les équilibres sociaux locaux, au point d'expliquer une part de la violence politique en Corse (Crettiez, 1998 ; Martinetti, 2007). De fait, les Etats et les régions concernées tentent de quantifier et comprendre le phénomène afin, le cas échéant, d'en réguler les effets les plus pervers. Or, la quantification du phénomène touristique n'est pas simple. Il est possible de comptabiliser un certain nombre d'offres touristiques. Il est possible de produire des statistiques de fréquentations dès lors qu'il y a une organisation suffisante (paiement d'un droit d'accès, réservation d'un service...). Il est en revanche extrêmement compliqué de comprendre et de localiser les effets de notoriété qui motivent le tourisme et en déterminent la distribution spatiale. A titre d'exemple, un visiteur peut motiver sa venue en Provence par le mythe des calanques de Marseille (où il n'y a aucun droit d'accès, ni moyen de comptabiliser la fréquentation) et louer son hébergement dans un village à 50 km. Dans ce cas, ce n'est pas l'offre d'hébergement du village qui a fait le tourisme, mais un élément subjectif non comptabilisé (le paysage des calanques). De la même façon, un touriste peut être attiré par les paysages sahariens et loger à Djerba. Nous faisons l'hypothèse que les éléments subjectifs non comptabilisables sont des ressorts de première importance du tourisme de masse (les paysages typiques, le mythe des plages ensoleillées, le dépaysement...). La régulation des influences du tourisme de masse sur les zones difficiles nécessite donc de disposer de méthodes d'analyse des éléments subjectifs construisant les notoriétés touristiques et de localisation de leurs effets. En effet dans ces régions là particulièrement, le décalage entre zone attractive et zone d'accueil effectif crée des déséquilibres sociaux (concentration des revenus du tourisme) d'autant plus importants que les revenus non touristiques sont faibles.

Dans cette perspective, nous développons une proposition méthodologique, avec la Corse choisie comme cas d'étude. Dans cette île, comme dans la Sardaigne voisine, un tourisme littoral très concentré engendre des conflits liés à la volonté de préserver les sites remarquables (Fiori *et al.*, 2008) alors même qu'une grande partie de l'attractivité de l'île semble liée à l'intérieur. Nous présentons tout d'abord la difficulté à obtenir des statistiques suffisantes pour refléter une dynamique touristique localisée. Nous développons ensuite la

¹ INRA, Ecodéveloppement, UR 0767, Avignon

² Idem, Plateforme Amandes.Txt, SAE2-SAD

³ INRA, LRDE, Corte. Nous tenons également à remercier Jean-Yves Gambotti et Annet Bruins pour leur aide technique.

justification et le protocole d'analyse de la notoriété touristique par l'analyse textuelle. Puis nous donnons quelques résultats du test et de la mise en œuvre de la méthodologie pour conclure par un inventaire des perspectives possibles à partir de ce premier test méthodologique.

I De la production de statistiques sur des phénomènes flous.

Quel est l'intérêt des méthodes d'analyse textuelle pour élaborer des variables explicatives de processus spatiaux faiblement couverts par les dispositifs statistiques classiques ? "*Pour compter ou mesurer des objets, il faut les identifier et les qualifier au préalable, ou définir des étalons de mesure à partir de propriétés qualitatives*" (Desrosières, 1995). Il s'agit donc d'une part de qualifier en établissant des classes d'équivalences rassemblant des entités que l'on juge proches par convention et d'autre part de mesurer l'importance numérique de ces classes. Or, la lourdeur des *investissements de forme* (Thévenot, 1985), qu'il s'agisse d'équipements, de personnels ou de conventions, explique que certains domaines, espaces ou secteurs d'activités, enjeux forts de gouvernement sont bien couverts (population, échanges commerciaux...), alors que d'autres sont décrits plus faiblement ou plus tardivement (environnement, discriminations...). La notoriété fait partie des domaines d'analyse nécessitant de lourds investissements de forme et ne répondant pas à un enjeu fort pour un Etat. Son analyse nécessite donc de s'interroger sur la construction d'indicateurs statistiques reflétant des phénomènes complexes faiblement équipés statistiquement.

Dans cette perspective, nous proposons de recourir à l'analyse textuelle informatisée pour mettre sous forme de variables quantifiables des textes reflétant la notoriété. Nous n'entrerons pas dans le détail des familles de méthodes relevant de ce vaste champ méthodologique (voir: Demazière *et al.*, 2006 ; Jenny, 1996 ; Roberts, 1997). De manière générale, il s'agit de partir d'une collection de texte (corpus) construite pour représenter de manière pertinente le phénomène à étudier, afin d'en extraire des propriétés à partir d'opérations de dénomination (ou codage) et de comptage de ces codes. Ces outils permettent l'établissement d'un lien entre les opérations d'analyse quantitative et qualitative en dépassant un clivage stérile, pourtant structurant les sciences sociales, légitimant de manière exclusive l'une ou l'autre des approches (Roberts, 1997). Les logiciels d'analyse textuelle sont des opérateurs de mise en équivalence entre le texte en langage naturel et un langage statistique normalisé et quantifié. L'interprétation des données s'appuie sur le postulat, largement partagé, d'une valeur sociale de la répétition des codes (Reinert, 2003). Ainsi, l'occurrence d'un mot ou d'un thème dans le discours reflète la valeur sociale que le locuteur lui attribue.

II La mesure de la notoriété touristique des communes corses à partir d'un corpus de guides touristiques

A. Statistiques et analyse du tourisme en tant que phénomène spatial

1. Le déficit de la description statistique du phénomène touristique en France.

En France, différents opérateurs (INSEE et le Ministère du Tourisme en particulier) produisent des statistiques censées permettre l'estimation de la dynamique territoriale du point de vue du tourisme. Or, ces informations sont réparties dans plusieurs bases de données dont la production suit des règles hétérogènes (méthodes, périodicité, unités spatiales) (Tableau 4). Ainsi, malgré la position stratégique du tourisme dans l'économie française, c'est un secteur dont la réalité est faiblement renseignée (Céron et Dubois, 2000), notamment du fait d'approximations complexes à lever, sauf à "bagueer les touristes" (Castille, 1998). Cette même faiblesse de renseignement existe dans le rapport entre tourisme et environnement (Dubois et Rechatin, 2000). Une première réponse est de mettre en place des dispositifs

d'enquête destinés à suivre en permanence les flux aux frontières et les déplacements touristiques des nationaux⁴ (Dreyfus-Signoles, 2002). Or, construire des dispositifs d'enquêtes suffisamment précis sur des phénomènes flous, nécessite de tels moyens qu'à notre connaissance aucun Etat ne l'a mis en œuvre. Par exemple, le bureau de la statistique de la Direction du Tourisme en France ne comptait que 6 employés en 2003 (Ceron et Dubois, 2003). Du côté du monde académique, il existe un consensus pour considérer que le matériel statistique existant est inadéquat à la correcte compréhension de l'objet touristique (Dewailly, 2006) et que les définitions "officielles" sont floues⁵ (Urbain, 1998). Des controverses font d'ailleurs suite à ce consensus et empêchent de parvenir à des définitions satisfaisantes du phénomène touristique (Lefort, 2007).

Le phénomène touristique est en effet complexe par son caractère mouvant dans l'espace (mobilité des touristes) et dans le temps (saisonnalités des activités). Par ailleurs, de nombreuses pratiques sociales sont associées au tourisme, des opérateurs professionnels aux pratiques individuelles des touristes, et suivent des modalités très diverses, en particulier du point de vue des mobilités (tourisme voyage, vacances, villégiature, excursion). La construction des indicateurs touristiques spatialisés (Tableau 4) relève donc de l'utilisation de données secondaires issues de divers secteurs statistiques non ciblés sur le tourisme (Dubois, 2005). Outre le détournement par rapport à des objectifs initiaux (par ex sur l'estimation de l'emploi touristique), la méthode a des limites en particulier concernant l'adéquation des échelles spatiales (parfois non exhaustives, trop englobantes ou trop restreintes) et temporelles (périodicité variables) des données. In fine, la description du phénomène touristique est principalement le fait de données sur les capacités d'accueil, la fréquentation et marginalement sur les valeurs et pratiques sociales des touristes associées aux lieux (Moisy, 2001). Or les 2 deux premiers domaines ne suffisent qu'à informer partiellement les dynamiques spatiales liées au tourisme. Il existe bien des disjonctions entre ces différentes dimensions, une capacité d'accueil existe mais peut être en sous ou en sur-utilisation (Mao, 2003) de même un lieu peut assurer l'hébergement effectif ou posséder une forte proportion d'emploi touristique mais n'est n'être pas le lieu de destination des visites (Violier, 2007) ou de forte valeur sociale (Mao, 2003).

2. Notoriété : une nouvelle dimension pour cerner la spatialisation du tourisme ?

De nombreux auteurs, en particulier géographes, sans nier l'importance de l'espace matériel (qualités physiques, équipement, accessibilité des lieux), insistent sur le rôle de l'image, de la représentation et des mythes attachés aux lieux touristiques⁶ (Cazes, 1992 ; Flament, 1975 ; Miossec, 1977). L'image de cet espace est médiatisée et assure ainsi le maintien de l'activité touristique et de l'attractivité d'un lieu (Tritz, 2001, 2003, 2006). Il ne suffit pas qu'un lieu existe, il est nécessaire que le touriste le sache pour accroître la probabilité qu'il le fréquente. Il y a donc un lien entre l'image perçue par un touriste et la fréquentation du lieu non superposable au lieu de résidence (Mao, 2003 ; Violier, 2007). La notoriété est cette valeur du lieu connue de tous ou du moins du plus grand nombre. Il faudrait d'ailleurs dire les valeurs puisqu'un lieu touristique est reconnu pour ses différentes valeurs sociales progressivement attachées à l'univers du tourisme : dimension esthétique, culturelle, récréative, sensuelle...

⁴ Depuis 1991, au sein d'un Observatoire National du Tourisme.

⁵ Données très tôt (1937) par l'Office Mondial du Tourisme, la définition d'un touriste est "toute personne en dehors de sa résidence habituelle pour une durée d'au moins 24 heures (ou nuit) et de 4 mois au plus pour l'un des motifs suivants : agrément (vacances et séjours de fin de semaine) ; santé (thermalisme, thalassothérapie) ; missions ou réunions de toutes sortes (congrès, séminaires, pèlerinages, manifestations sportives...) ; voyages d'affaires, déplacements professionnels ; voyages scolaires" Urbain J.D., 1998. L'idiot du voyage, histoires de touristes. Payot et Rivages, 430 p.

⁶ Miossec (1977) : "l'espace touristique c'est avant tout une image" ; Flament (1975) : "le succès et le développement d'une station dépendent très largement de l'idée que s'en fait la société. Une portion d'espace ne devient touristique que si elle coïncide avec ses goûts, ses mythes".

(Tritz, 2006). Ces valeurs des lieux sont diffusés par différents médias (brochures, revues spécialisées, sites web...) parmi lesquelles le guide touristique tient une position particulière. Nous allons montrer comment la littérature dans sa diversité (Daye, 2007) converge pour affirmer cette place prépondérante des guides touristiques dans la formation de la notoriété touristique des lieux.

B. Saisir la notoriété des territoires par l'analyse textuelle des guides touristiques

1. Justifications théoriques et empiriques de la pertinence des guides touristiques comme corpus

Un état de la littérature prenant comme objet les guides touristiques, en particulier les "*Tourism Studies*" mais aussi les approches géographiques et littéraires en France, permet de justifier que cet objet littéraire soit utilisé dans une perspective métrologique.

Le guide touristique est un texte appréhendé de deux façons dans les recherches : d'une part, comme un reflet de la valeur des lieux telle qu'elle est présente dans une culture donnée, d'autre part par son caractère fortement performatif intervenant donc dans les comportements et représentations des touristes. Ces deux points de vue sont des arguments forts de l'utilisation des guides dans la construction d'indicateurs de notoriété.

a) Guide reflet d'une représentation sédimentée de la valeur des lieux

Le guides est un matériau fréquemment mobilisé dans l'analyse de l'histoire des mobilités ou des identités nationales (Koshar, 1998), de la construction des lieux et de leurs représentations (Bonin, 2001 ; Santos, 2004 ; Saunier, 1993 ; Toulhier, 2000 ; Tritz, 2003 ; Young, 1999). Le texte est un marqueur de la valeur sociale d'un lieu (Richards, 2002), de sa notoriété ou des mythes qu'il génère (Barthes, 1957). L'analyse historique de la construction de ce genre littéraire met en évidence sa propension à véhiculer des clichés et représentations socialement partagés et diffusés (Chabaud *et al.*, 2000 ; Dann, 2001 ; Guilcher, 2004). Le genre s'ancre dans la tradition du récit de voyage mais s'est mué en un discours utilitaire, voire promotionnel et prescripteur, intégré dans l'industrie touristique (Gritti, 1967 ; Guilcher, 2000 ; Kerbrat-Oriccionni, 2005 ; Saunier, 1993). Une des caractéristiques de la construction des guides est la méthode minimaliste de leur construction, consistant à reprendre principalement des informations déjà existantes sur la nature des lieux (littérature de voyage, guides concurrents, versions antérieures, histoire et géographie locale) et à ne conduire que marginalement des enquêtes (Bonin, 2001). Ce phénomène de reprise se retrouve également dans les traductions. Ainsi, "la très grande majorité des guides disponibles en Italie et en Espagne sont en fait des traductions du Guide Michelin, du Guide Vert, ou d'autres guides" (Tritz, 2006). Même s'il ne s'agit pas de traductions fidèles, Tritz, en comparant des guides français et anglais de la Côte d'Azur différents (Michelin / Insider's Guide), montre la très faible différence dans les modalités de médiatisation des lieux (Tritz, 2001).

Par son objectif et son mode de construction, le genre relève donc d'un discours universel maximisant le consensus sur les lieux de notoriété. Ainsi, il existe une hiérarchie stable des lieux à visiter avec néanmoins l'intégration de nouveautés, par exemple intégration du tourisme intérieur sur la Côte d'Azur (Tritz, 2003), ou la modification du mode d'organisation des visites (Bonneau, Violer, 2000).

b) Guide performateur des représentations et des comportements du touriste

Différentes approches théoriques et empiriques, les unes critiques (Barthes, 1957 ; Dann, 1996), les autres ethnométhodologiques ou marketing et optimisatrices (Butler, 1990 ;

Gartner, 1993 ; Hunt, 1975), mettent en évidence les effets du texte sur le touriste et son expérience de la visite (Daye, 2007). Dans la première perspective, le guide est un instrument d'aliénation, de contrôle social ou d'aveuglement des touristes limitant les potentialités du touriste (Cassou, 1967 ; Dann, 1996 ; MacGregor, 2000 ; Zillinger, 2006). Dans la seconde perspective, le touriste est doté de compétences d'interprétation et construit son expérience avec le guide (Brown, 2007 ; MacCannell, 2001 ; Mercille, 2005 ; Urbain, 1998 ; Urry, 1990). Ce n'est donc pas l'existence d'un effet des guides sur les touristes mais la force de cet effet et l'autonomie du touriste qui sont en jeu.

Premièrement, le guide influence la perception qu'a le touriste des lieux qu'il visite, le "tourist gaze" (MacGregor, 2000 ; Perkins et Thorns, 2001 ; Urry, 1990, 1995), la construction de l'image des lieux (Santos, 2004). Des études empiriques montrent ainsi une convergence entre les images délivrées par les guides et les images individuelles postérieures à la visite (MacGregor, 2000 ; Mercille, 2005). Deuxièmement, l'image des lieux influe in fine sur les choix individuels, en particulier de mobilité (Jenkins, 1999 ; Prod'homme, 2002 ; Violier, 2007 ; Zillinger, 2006). MacGregor (2000) montre que les touristes sont à la recherche des éléments dont ils ont une connaissance, ne serait-ce que partielle (le connu et l'imaginé), les autres aspects des lieux restant dans la sphère de l'inconnu ou de l'invisible. S'agissant de l'effet sur les mobilités, les auteurs mettent en avant le haut degré d'incertitude spatiale, temporelle et sociale que constitue le voyage (Brown, 2007 ; Ladwein, 2002 ; Violier, 2007). Dès lors le touriste s'équipe du guide, véritable "prothèse visuelle" étoffant "symboliquement l'expérience visuelle" (Ladwein, 2002), pour positionner les lieux dans l'espace et le temps, en évitant ainsi le risque d'errance et d'insignifiance. Le guide est également un objet pratique qui par son format, la nature pratique, la structuration et la normalisation de son contenu est intégrable dans l'activité quotidienne du touriste au prix d'un travail de contextualisation, de combinaison avec d'autres informations (cartes, conversations...) et d'interactions sociales (Brown, 2007).

c) L'importance économique et pratique des guides

Au-delà de ces considérations théoriques, quelques données empiriques et statistiques permettent de se convaincre de l'importance qu'ont les guides avant et au cours de l'expérience des touristes. Quinlan (2005) passe en revue une série d'articles qui montrent le pourcentage important, variable cependant, de touristes utilisant un guide touristique (92 % des touristes à Tana Toraja en Indonésie (MacGregor, 2000), 65 % chez les Japonais visitant l'Alaska (Milner *et al.*, 2000)... En France, les guides occupent une position importante dans l'information du touriste, ainsi environ 68 % des 40 millions de Français qui partent en vacances au moins une fois par an achètent des guides de tourisme (Livre Hebdo, 2007). 8,6 millions de guides ont été vendus en 2005 en France (Petit, 2007).

2. Dispositif opérationnel pour formaliser statistiquement la valeur des lieux représentée dans les guides

a) Les méthodologies existantes

Des méthodes d'analyse textuelle des guides visant la production d'indicateurs numérique de notoriété ont déjà été utilisées dans la littérature géographique mais elles nécessitent la construction de base de données alimentées par une analyse manuelle du contenu des guides. Par exemple, des cartes de densité au km² d'étoiles sont produites à partir de la géolocalisation des sites étoilés du Guide Vert (Knafou, 1997 ; Mao, 2003). Tritz (2001, 2006) construit de manière manuelle une base de données à partir des lieux cités (classés en haut-lieux et lieux secondaires) dans les annexes des guides (GR, GV et Guide Bleu) à l'échelle du bassin nord méditerranéen. Elle représente les résultats sous forme raster pour les traiter à partir de méthodes d'analyse d'image afin de modéliser la répartition spatiale des lieux touristiques

(modèle centre-périphérie). L'analyse de contenu peut également porter sur le décompte de l'iconographie des guides (Tritz, 2006) ou des photos géoréférencées sur le web (Girardin *et al.*, 2008).

Notre objectif étant de tester une approche reproductible automatisant le plus possible la mesure en particulier pour une extension éventuelle vers l'analyse du web, nous avons adopté une autre méthodologie. Il s'agit ici de présenter la méthodologie et les résultats du test.

b) Méthodologie semi-automatisée de mesure de la notoriété

Toute analyse textuelle commence par la définition d'un corpus que l'on juge représentatif du phénomène étudié. Le marché des guides est couvert par une multitude de produits qui adoptent des politiques éditoriales contrastées pour viser des niches différentes. Par exemple, le Guide Bleu par son coût et son contenu culturel cible un public aisé alors que le Guide du Routard vise un public populaire et/ou jeune (Chabaud *et al.*, 2000). Nous avons choisi d'analyser les principaux guides généralistes couvrant la majorité de la part de marché (Leboucher, 2007 ; Livre Hebdo, 2007 ; Petit, 2007) et les guides servant de références dans les traductions. Il s'agit donc d'analyser les guides ayant potentiellement le plus d'impact sur la formation de la notoriété dans la population. Dans le cas de la Corse, nous avons utilisé la référence historique, le Guide Vert (GV), avec 20 % de part de marché et le renouveau, le Guide du Routard (GR) avec 40 % de part de marché. Nous avons choisi d'ajouter le Lonely Planet (LP) qui est un guide d'un éditeur australien en cours d'implantation sur le marché français. Ces différents guides ont été numérisés par reconnaissance de caractère. Sans détailler le protocole, il est important de noter que cette opération nécessite un important contrôle de la qualité des textes produits.

La seconde opération consiste à mettre en équivalence ce corpus en langage naturel et le système statistique de référence par le biais d'une analyse textuelle semi-automatisée. Ceci est permis car la commune est majoritairement l'entité géographique de référence dans les guides et dans le système INSEE.

L'outil Prospéro, développé initialement pour l'analyse des dossiers complexes en sociologie (Chateauraynaud, 2003), permet de produire une analyse qualitative fine, grâce à de multiples possibilités de codage, de création d'expressions et de construction des réseaux lexicaux de termes cibles. Son principal intérêt dans cette méthodologie réside dans sa capacité à contourner une difficulté de la langue liée à la multiplicité des façons de désigner une entité, dans notre cas plus particulièrement les multiples graphies des noms de communes. Le logiciel permet ainsi de regrouper les termes désignant le même territoire sous une même entité. Par exemple, Sainte Lucie de Tallano, SAINTE LUCIE DE TALLANO, Ste-Lucie-de-Tallano... renvoient tous à la même entité qui est désignée dans le référentiel statistique de l'INSEE par l'identifiant 2A308. Ce travail de mise en équivalence de l'hétérogénéité du langage naturel et des catégories statistiques normalisées a été automatisé à l'aide d'une macro sous Excel. Nous avons appliqué des règles de variation des graphies (Majuscule, traits d'union, raccourcis...), préalablement inventoriées sur un tableau du référentiel statistique (354 communes corses : 2118 graphies). A partir de ce fichier établissant la correspondance entre l'identifiant statistique et les graphies des communes, Prospéro extrait un tableau de comptage issu de l'analyse du corpus. L'indicateur de notoriété brut est donc l'occurrence de citation des noms de communes dans les guides.

III Résultats du test méthodologique

A. Des ambiguïtés linguistiques et statistiques ne permettent pas une automatisation complète de la construction des indicateurs

Lors de la mise en œuvre de la méthodologie nous avons été confrontés aux effets de l'évolution des découpages administratifs et statistiques reflétant la complexité de l'organisation de l'espace géographique et administratif corse. En effet, la géographie corse se caractérise par une opposition sur une courte distance entre un espace littoral et un espace montagnard ainsi que la présence de nombreux hameaux dans les communes. Afin de refléter au mieux cette complexité géographique et suite à des évolutions démographiques et à la départementalisation, des fusions (par ex Mezzavia fusionné avec Ajaccio) puis des redécoupages de communes ont été opérés par l'Administration et/ou l'INSEE. L'INSEE, dans son travail statistique, utilise ainsi la notion de "communes éclatées" pour distinguer le littoral de la montagne revenant ainsi sur certaines fusions administratives (par ex: Porto, station balnéaire hameau de Ota, est individualisé comme "Porto") ou générant de nouveaux découpages (par ex Molini issu d'Albitreccia).

Ce problème se résout en établissant une table de correspondance entre ces différents découpages statistiques. Mais les références toponymiques utilisées par chaque guide se sont avérées contrastées, ce qui tend à générer du bruit dans le comptage des citations. En effet, le recours au fractionnement de certaines communes, entre hameaux et commune mère, privilégie ces hameaux alors que les hameaux des autres communes ne sont pas pris en compte. Il serait donc nécessaire de compléter les fichiers d'équivalence avec l'ensemble des hameaux de toutes les communes. Ceci étant, cette opération pourrait induire des risques de générer un autre bruit lié à l'homonymie des toponymes.

L'homonymie est d'ailleurs une des principales sources d'ambiguïté dans le travail sur les toponymes. Nous l'avons peu rencontrée en Corse mais le cas se présente en région PACA (24/964 communes vrais doublons) et encore plus fortement à l'échelle française (418 vrais doublons et 2536 quasi-doublons sur 36783 communes).

Après examen, à la variation des graphies s'ajoutent les fautes d'orthographe présentes dans les guides ou issues du travail de numérisation des textes mais aussi des variations dues à la francisation des noms corses (par ex: les doubles lettres corses (gg, cc...) sont simplifiées en français).

Enfin, un problème important apparaît autour des toponymes à intégrer ou exclure du décompte communal. Nous avons déjà évoqué la question de l'insertion ou non des noms de hameaux. Nous avons également vu apparaître des ambiguïtés liées à l'existence de toponymes comprenant des noms de communes. Ainsi, les Monte d'Oro, Monte Estremo etc... sont comptés sous la commune Monte si l'on ne prend pas garde à "figer" ces expressions⁷. Mais par ailleurs, le toponyme forêt de Vizzavona qui se trouve effectivement majoritairement sur l'ancienne commune de Vizzavona (Vivario) n'a pas été figée. Ainsi, on ne peut pas traiter cette question avec un gel massif des expressions mais avec un codage ciblé nécessitant un important travail de vérification.

B. Signification des indicateurs produits

1. Comparaison des indices de notoriété des différents guides

Globalement, il y a une très bonne convergence des manières de refléter la notoriété entre les différents guides. Cela se lit par la lecture des cartes (Figure 1) et par les corrélations entre nombre de citations ou des correspondances entre communes cités dans chaque guide (Tableau 1). Un patron spatial ressort de la cartographie des différents indicateurs. Les zones

⁷ Le Monte dans l'expression figée de "Monte d'Oro" n'est plus comptabilisée sous Monte. La construction des expressions est semi-automatisée dans Prospéro

de plus forte notoriété se situent sur le littoral aux alentours de Calvi au Nord-ouest et aux alentours de Bonifacio et Porto-Vecchio au Sud en passant par la ligne de crête centrale de l'île. On note donc une relativement faible mise en exergue des communes littorales et de certaines territoires pourtant connus par ailleurs pour leur attractivité. C'est en particulier le cas des communes du Cap Corse qui ne ressortent que faiblement. Or les occurrences des désignations de ce territoire dans les guides analysés sont loin d'être négligeables (281 occurrences sur les 3 guides juste derrière Corte). Ce constat pose le problème de la mesure de notoriété sur les entités territoriales supra-communales à forte identité qui n'a pas été intégrée dans la méthodologie à ce stade de développement. Mais, il est envisageable d'intégrer ces territoires, identifiés dans les index des guides, au fichier d'équivalence de Prospéro et de mettre en œuvre une procédure répartissant leurs notoriétés sur l'ensemble des communes les composant.

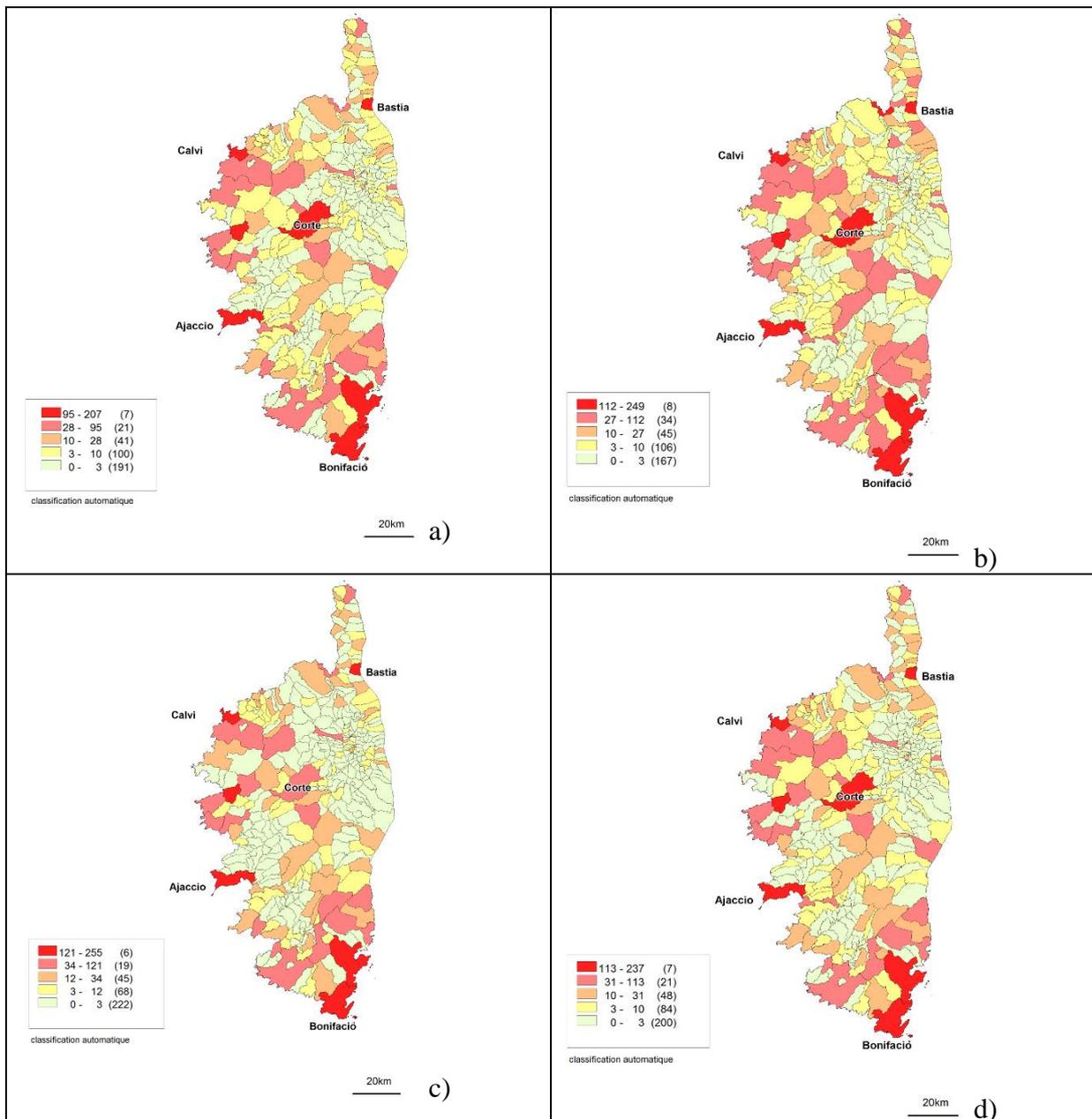


Figure 1 : Indices de notoriété brut du GR (a), du GV (b) du LP (c) et moyenne des trois guides (d)

Tableau 1 : corrélations entre les occurrences brutes des noms (a) et taux de correspondances (b) des communes de chaque guide en Corse .

a)	GR	GV	LP
GR	1		
GV	0.967	1	
LP	0.97	0.97	1

b)	GR	GV	LP
GR	1		
GV	0.66	1	
LP	0.71	0.64	1

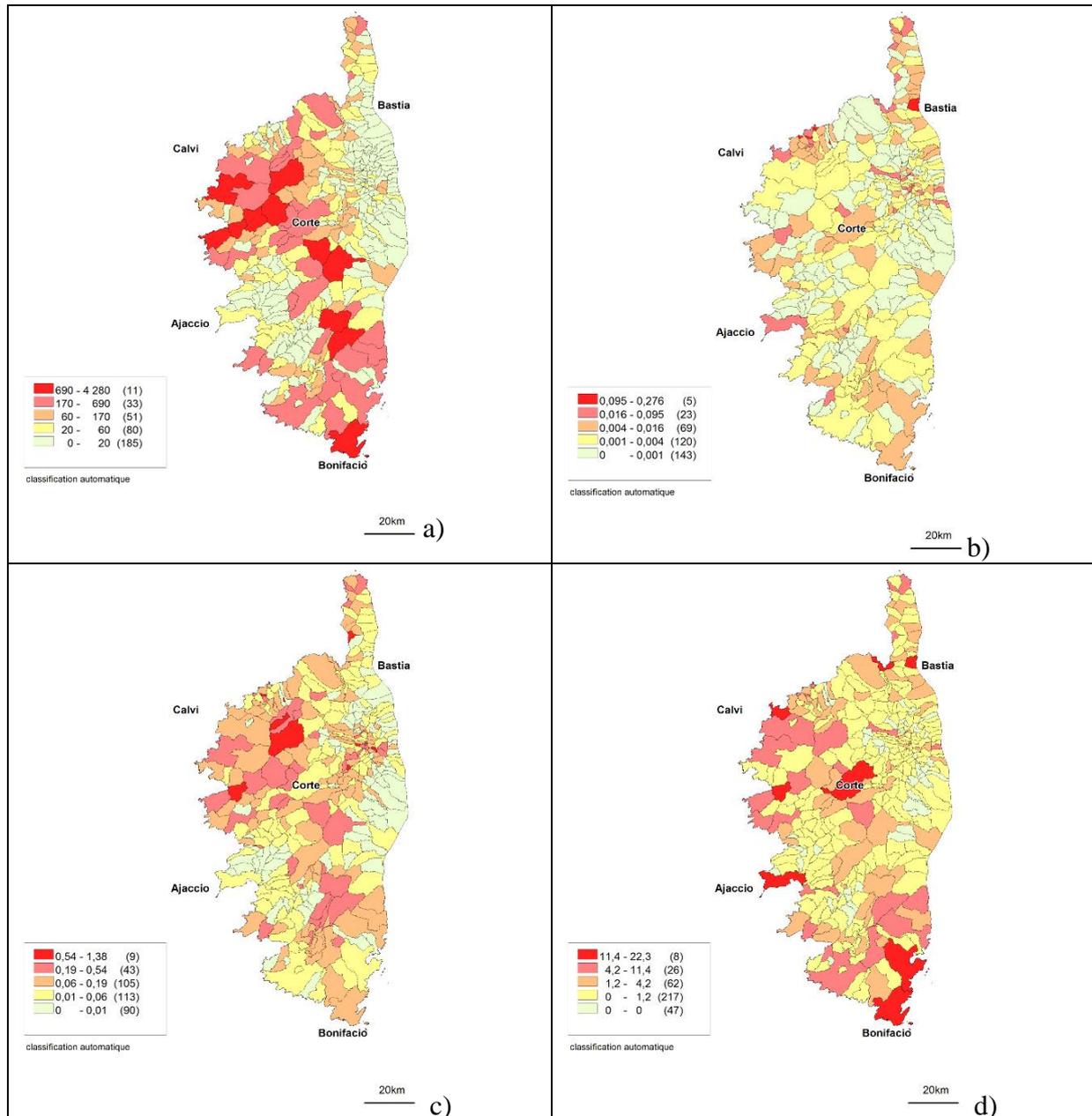


Figure 2 : Indices de notoriété moyens pondérés par la densité de population (a), la surface communale (b) la population (c) et le logarithme de la population (d)

Par ailleurs, les données brutes mettent en exergue les villes ce qui invite à pondérer les indicateurs (Figure 2). Les différentes pondérations mises en œuvre à ce stade (population, densité de population) ne remettent pas en cause ce patron spatial montrant ainsi la validité de la méthode de mesure de l'indicateur de notoriété. Seule la pondération par la surface communale affecte fortement le patron spatial de la notoriété mais ceci reste difficilement interprétable, la distribution des tailles de communes étant peu rattachable à un phénomène identifiable sous-jacent contrairement à la population et la densité de population qui sont liés

aux gradients urbain-rural. Peu peuplés et étendus, comme tous les villages forestiers, les villages de la crête centrale de l'île ont une notoriété par habitant très élevée. Entre plages et forêts facilement accessibles par la route, se dessine en creux une carte de la "non notoriété", qui regroupe une majorité des villages corses situés en moyenne montagne. Seuls échappent à cette non-notoriété les villages de moyenne montagne bénéficiant d'une grande originalité culturelle et paysagère, par exemple des villages de la Castagniccia au nord-est, par pailleurs peu peuplés et peu étendus.

Dans le détail qualitatif, et de manière marginale, la mesure de la notoriété reflète quelques différences liées à des variations de politiques éditoriales des guides. La très grande majorité des communes est citée par les différents guides mais le GV, suivi de près du GR, couvre plus largement le territoire corse que le LP (Tableau 2). Le GV est par ailleurs le plus prolixe en citations de communes, et donc en détails descriptifs, alors que le GR fait moins références aux communes. Un noyau de 169 communes est décrit par tous les guides (48% des communes corses soit 93 % du total des occurrences) mais chaque guide met en avant des communes qui lui sont spécifiques, c'est particulièrement le cas encore une fois du GV.

Tableau 2 : Couverture géographique des différents guides

Guides	Nbre communes	Som. citations
Guide du routard	255	2985
Guide Vert	273	4175
Lonely Planet	226	3492
Communes couverte par les 3 guides	336	10652
Communes en commun (3 guides)	169	9882
Communes en commun (2 guides)	81	553
Communes spécifiques	86	217
Communes de Corse	354	-

Tableau 3 : caractéristiques des communes spécifiques de chaque guide

	GR	GV	LP
nb de communes spécifiques	26	45	15
Somme des occurrences pour les communes spécifiques	56	133	28
Citations/commune	2.15	2.96	1.87
Moy. hab	159.4	138.4	120.1
Moy. Résid. 2nd	57.2	54.95	76.35
Moy. empl camping 07	0	0.73	40.07
Moy chambres hôtel 07	3	2.09	1.2
Moy. Capacité gîte rural	6.35	1.36	1.64
Moy. nb emploi tourist.	1.45	1.36	1.64
% cit Attrac Mer	8.3	10.2	16.1
% cit Attrac Montagne	31.3	34.7	38.7
% cit Attrac Forêt	29.2	19.4	22.6
% cit Attrac Campagne	31.3	35.7	22.6

De manière générale, les communes spécifiques sont peu nombreuses et peu citées. Nous avons croisé les communes spécifiques avec quelques indicateurs démographiques et touristiques afin de caractériser les communes mises en avant par chaque guide. Globalement ces communes sont peu peuplées. Les communes spécifiques au LP sont plus touristiques (présence de résidences secondaires et campings et emploi touristique plus important en moyenne), ce guide met le plus en avant les communes dont une des causes d'attraction est la mer même si la montagne apparaît majoritaire (Tableau 3). Le GR cite spécifiquement des communes plus peuplées en moyenne avec une capacité d'accueil hôtelier et surtout en gîtes ruraux plus importante. Ces communes sont plutôt attractives pour leur aspect montagnard et forestier, on pense ici aux communes centrales autour du Chemin de Grand Randonnée (GR20). Le GV se distingue par sa description dense des communes qu'il cite spécifiquement qui apparaissent plutôt plus attractives pour leur aspect campagnard et moins pour l'aspect forestier.

Nous avons utilisé les possibilités d'analyse qualitative de Prospéro pour rendre compte de la différence de traitement des différents guides en ciblant les communes spécifiques et les communes avec les plus grands écarts de citations. L'analyse des mots spécifiques des réseaux lexicaux des communes spécifiques est assez significative de l'orientation éditoriale des

guides. Le GR est tourné vers l'arrière pays, ses gîtes d'étapes, son aspect montagnard, son relief (mots spécifiques: *crête, panorama, beauté, promontoire, défilés, hauteurs*) et les balades que cela permet comme le montrait le patron spatial précédemment décrit. Le GV est davantage orienté vers les activités pratiquées (*thalassothérapie, thermalisme*) en particulier dans le milieu rural (*agriculture, élevage, mine, artisanat...*) dans les communes qu'il met spécifiquement en avant. Il insiste donc sur les produits (*vin, bière, brocciu, confiserie...*) et le patrimoine rural (*chapelle, calvaire, murailles*). Il développe enfin une dimension culturelle (*art, fresques, oeuvres, identité, dialecte*) et historique (*Bonaparte*). Enfin le LP met l'accent dans ses communes spécifiques sur les activités sportives et de loisir (*cheval, vélo, surf, funboard*) plutôt liée au littoral (*mérus, éponges, tortues...*) mais aussi sur le patrimoine et les occupations urbaines (*musée, esplanade, cathédrales, café, festivités, télévision*).

Ces orientations se retrouvent en partie dans les différences de traitement des communes avec les plus grands écarts de citation. Ainsi, le GR et le GV citent moins les grandes villes littorales comparativement au LP (respectivement 23 % des citations et 22.9 % pour les 5 premières villes en population contre 27 % pour le LP). Pour les autres communes à fort écart de citation, les descriptions sont contrastées. Le LP a tendance à surreprésenter les villes et communes littorales en mettant en avant leur aspect balnéaire. Pour ce type de commune, le GV donne une orientation historique plus prononcée y compris dans la Saint-Tropez corse (St-Florent).

Pour conclure, la hiérarchie des valeurs de lieux sur l'ensemble de la Corse est convergente entre les différents guides, il y a donc bien une forte structuration du phénomène de notoriété reflété d'ailleurs par une distribution en loi de puissance. Mais, dans le détail, leurs politiques éditoriales, et les niches de touristes visées, orientent les types de valeurs et les communes correspondantes mises en avant. Le GV et le GR semblent être les plus informatifs du point de vue géographique et descriptif.

2. La mesure de la notoriété comparée à quelques indicateurs touristiques.

En comparant la distribution spatiale des indicateurs de notoriété avec les autres indicateurs touristiques, nous cherchons à mieux cerner l'apport de la mesure de la notoriété sur la compréhension du phénomène touristique (Figure 3). La répartition de la population résidente suit un tropisme lié au littoral et à l'axe routier entre Ajaccio et Corte. Ce tropisme se retrouve également dans la distribution spatiale des emplacements de camping. Par contre, la densité de chambres d'hôtel par habitant suit une répartition qui se rapproche de celle de la notoriété. Les données sur les résidences secondaires sont reconnues comme étant peu fiables en Corse car elles mêlent à la fois la résidence touristique et les "maisons de famille au village". Leur patron spatial se distingue d'ailleurs de celui des accueils strictement touristiques (camping et hôtels) ce qui indique que ce phénomène suit une logique propre.

N'y aurait il pas dans ces différents constats l'indice d'un phénomène de distinction sociale entre un tourisme populaire et balnéaire (campings littoraux) peu socialement valorisé, sans être nié, dans les guides touristiques et un tourisme plus élitiste (hôtels), "un tourisme de la profondeur" (Urbain, 1998), associé à des valeurs socialement légitimées prônées par les guides touristiques (la culture, l'effort, l'exploration, le beau, la profondeur et l'authenticité... opposée à l'ignorance, la passivité et le laid, la superficialité et le factice...). Ainsi, Urbain (1998) montre l'existence d'un "mépris" social du touriste⁸. Cette représentation négative du tourisme de masse, relativement bien reflétée pas les indicateurs d'hébergement touristique,

⁸ "le touriste paraît avoir fait sien le mépris que le voyageur lui porte. Il l'a intériorisé. (...) le discours de promotion touristique doit être antitouristique" (p 125), à propos des publicités et des guides : "Tout en exacerbant la phobie du touriste, il s'agit de compenser son aversion antitouristique par une promesse de pérégrination singulière, extra, supra ou para touristique. Faire croire au touriste qu'il n'en est pas un" (p 127) (Urbain, 1998)

amène les guides, et plus largement les médias promotionnels, à mettre en avant des pratiques et des lieux socialement valorisés.

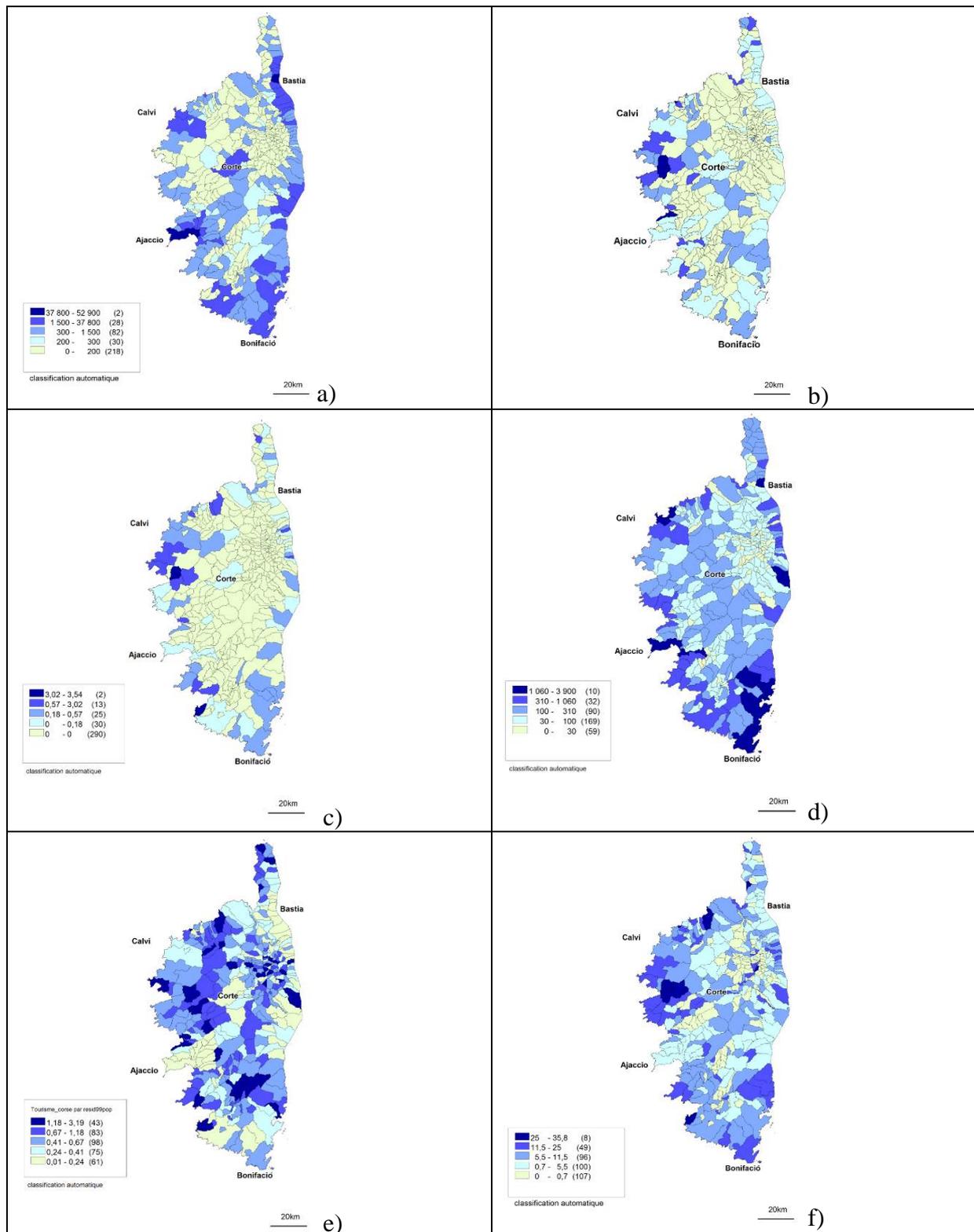


Figure 3 : Population communale 1999 (a) et quelques indices touristiques : Nombre de chambres d'hôtel par habitants (b), Nombre d'emplacements de camping par habitants (c), Nombre de résidences secondaires (d), résidences secondaires par habitants (e) et Part de l'emploi touristique (f)

Les indices de notoriété se différencient des indices touristiques de capacité d'accueil par la mise en avant des communes intérieures de l'arête centrale corse où se situent les paysages de reliefs et les activités de nature associées (balade et randonnée en particulier). Par l'interprétation de certains indicateurs touristiques (emplacements de camping et chambres d'hôtels), on retrouve deux pôles touristiques déjà reflétés par la notoriété : les alentours de Calvi au nord-ouest et les alentours de Bonifacio. Mais les indices touristiques tendent à mettre davantage en avant le littoral en particulier de la côte est.

L'indice de notoriété apporte donc bien une information autre. Ces résultats confirment l'importance de la distinction entre espace matériel et espace médiatisée que l'on trouve dans la littérature (Cazes, 1992 ; Flament, 1975 ; Miossec, 1977 ; Tritz, 2001). La confrontation des différents indices donne également des indications sur le modèle de mobilité touristique en Corse avec des pôles littoraux de résidentialité touristique à forte capacité d'accueil et forte notoriété, des "zones d'exploration" dans l'intérieur de l'île à population et capacité d'accueil limitée et enfin des zones de faible notoriété à forte population et capacité d'accueil (littoral est). Il se dessine ainsi le profil d'un touriste "dormant sur la côte et se promenant à l'intérieur".

IV Perspectives de développement de la méthodologie

La méthodologie que nous avons testée met en évidence la nécessité d'un contrôle humain de la chaîne de traitement s'appuyant sur une connaissance de l'espace géographique analysé. D'autres tests sont à mener pour mesurer l'effet de l'insertion des toponymes infra-communaux et supra-communaux dans le comptage des occurrences et du gel d'expressions toponymiques. Les ambiguïtés peuvent être contournées en corrigeant manuellement les données. L'outil Prospéro mobilisé pour ce test permet néanmoins d'enrichir l'analyse par une dimension qualitative que nous avons commencé à utiliser dans cette première contribution. Une analyse plus approfondie pourra permettre de qualifier davantage les sources de la notoriété touristique et classer les communes en fonction des grands types de valeurs (balnéaire, culturel, gastronomique, récréatif...). Cet approfondissement méthodologique est d'autant plus nécessaire pour mieux mesurer et expliquer les constats issus de cette première approche montrant que la notoriété touristique de cette île méditerranéenne est basée sur une attractivité balnéaire mais aussi sur la notoriété des sites intérieurs montagnards et forestiers et/ou à valeur culturelle. Cette bipolarisation de la notoriété de la Corse, déjà relevée dans les études marketing sur l'image globale de la Corse (KPMG, 2008), laisse une grande partie des communes intermédiaires "dans l'ombre" et ce d'autant plus que celles-ci ne sont pas situées sur des circuits routiers. Ceci tendrait à montrer que la notoriété touristique est très inégalement répartie sur le territoire y compris dans une île si densément fréquentée.

Dans une perspective d'élargissement à l'analyse de textes diffusés sur le Web, il apparaît nécessaire d'amender la méthodologie. Cet élargissement se justifie en effet car l'émergence d'internet tend à transformer la forme et les pratiques associées aux guides touristiques (Pachaud, 2006 ; Violier, 2007). L'accès au texte est grandement facilité à la fois pour le touriste mais aussi pour le chercheur qui analyse ses pratiques ce qui fait du web un corpus intéressant à analyser. L'élargissement peut également être géographique en passant d'analyses régionales à des analyses nationales voire internationales dès lors qu'un corpus et qu'un fond cartographique existent. L'élargissement peut enfin concerner les phénomènes que nous cherchons à mesurer dans le cadre de la thématique des marchés fonciers mais plus largement aussi dans les problématiques d'analyse des processus spatiaux. C'est par exemple le cas de la mesure des formes de conflictualité territoriale qui nécessitent la construction de bases de données textuelles importantes (presse, arrêt de tribunaux, procès-verbaux) (Kirat *et al.*, 2005 ; Melot *et al.*, 2008).

Etant donnée la taille du corpus que constitue le Web, il apparaît peu réaliste d'utiliser une méthode semi-automatisée. Par ailleurs, il serait nécessaire de construire un dispositif de fouille du web (crawler) permettant de filtrer l'information selon les sources, les emprises géographiques et thématiques. Prospéro, bien que très utile pour lever le problème de la multigraphie, ne permet pas de lever toutes les ambiguïtés associées à l'identité des noms de lieux que nous avons mis en évidence lors de notre test. D'autres méthodologies d'analyse textuelle développées dans les domaines du Traitement Automatique du Langage Naturel (TALN), du Web Sémantique et de l'Extraction des Connaissances sont davantage formalisées au plan linguistique et informatique et permettent de construire des raisonneurs qui désambigüisent les termes à partir de leur contexte. Sur le plan de la toponymie et de l'extraction de l'information spatiale, des travaux prometteurs ont déjà été identifiés (Lesbegueries, 2007 ; Schockaert *et al.*, 2008). Cette perspective nécessite un important travail de développement informatique et non plus le simple usage d'outils préexistants.

Au regard des questionnements que suscitent les dynamiques des zones difficiles, la méthodologie que nous avons mise en œuvre représente une voie d'analyse prometteuse. Il s'agit souvent de zones qui sont confrontées à la fois à des processus économiques et sociaux potentiellement perturbateurs dont les ressorts sont externes au territoire (le tourisme de masse, l'urbanisation des émigrés, l'agriculture industrielle...) et à la fois des zones souffrant d'un déficit d'informations centralisées rendant difficile l'étude des processus de changement et l'évaluation de leurs effets. La méthode que nous proposons permet, à partir de données facilement disponibles (les guides et des fonds cartographiques à des échelles peu précises), d'élaborer des indicateurs rendant compte de déterminants des tendances observées.

Bibliographie

2006. Loi 2006/437 du 14 avril 2006 portant diverses dispositions relatives au tourisme.
- Barthes R., 1957. Mythologies. Seuil, Paris.
- Bonin S., 2001. Paysages et représentations dans les guides touristiques. La Loire dans la collection des Guides-Joanne, Guides Bleus (1856 à nos jours). *L'espace géographique*, vol. 30, n° 2, pp. 11-126.
- Brown B., 2007. Working the problems of Tourism. *Annals of Tourism Research*, vol. 34, n° 2, pp. 364-383.
- Butler R., 1990. The Influence of the Media in Shaping International Tourist Patterns. *Tourism Recreation Research*, vol. 15, n° 2, pp. 46-53.
- Cassou J., 1967. Du Voyage au tourisme. *Communications*, n° 10, pp. 25-34.
- Castille D., 1998. L'emploi salarié touristique du Nord-Pas-de-Calais. *Profils*, n° 11, pp. 4.
- Cazes G., 1992. Fondements pour une géographie du tourisme et des loisirs., Bréal, 189 p.
- Ceron J.P. et Dubois G., 2003. Tourism and Sustainable Development Indicators: The Gap between Theoretical Demands and Practical Achievements. *Current Issues in Tourism*, vol. 6, n° 1.
- Céron J.P. et Dubois G., 2000. Les indicateurs du tourisme durable : un outil à manier avec discernement. *Cahiers Espaces*, n° 67, pp. 22.
- Ceron J.P., Dubois G., Armand L., Nadaud F. et Tardieu F., 2008. Utilisation des enquêtes sur la demande du tourisme national afin d'évaluer les émissions de gaz à effet de serre issues du transport de passagers longue distance. Forum International des Statistiques du Tourisme, Paris, 19-21 novembre 2008, OCDE, Ministère Français de l'Economie, de l'Industrie et de l'Emploi, http://www.forumstat.tourisme.gouv.fr/ftp/fr_S5_3_dubois.pdf.
- CERTU, 2007. Outils et méthodes pour l'observation urbaine. Inventaire communal: Ministère de l'Equipement, http://www.observation-urbaine.certu.equipement.gouv.fr/article.php?id_article=244.
- Chabaud G., Cohen E., Coquery N. et Penez J., 2000. Les guides imprimés du XVIe au XXe siècle. Villes, paysages, voyages. Belin, 704 p.
- Chateauraynaud F., 2003. Prospéro - Une technologie littéraire pour les sciences humaines. CNRS Editions, 406 p.
- Crettiez X., 1998. Lire la violence politique en Corse. *Les Cahiers de la sécurité intérieure*, n° 33, pp. 195-214.
- Dann G.S., 1996. The Language of Tourism. A Sociolinguistic Perspective. Wallingford: CAB International.
- Dann G.S., 2001. The Self-Admitted Use of Cliche, in the Language of Tourism. *Tourism Culture & Communication*, vol. 3, n° 1, pp. 1-14.
- Daye M., 2007. Developing a framework for the study of tourism and the media. Culture, Tourism and the Media, Nottingham, Nottingham University Business School, http://www.nottingham.ac.uk/ttri/pdf/2007_1.pdf.
- Demazière D., Brossaud C., Trabal P. et Van Meter K., 2006. Analyses textuelles en sociologie : Logiciels, méthodes, usages. PUR, 218 p.
- Desrosières A., 1995. Classer et mesurer. Les deux faces de l'argument statistique. *Réseaux*, n° 69.
- Dewailly J.M., 2006. Tourisme et géographie, entre pérégrinité et chaos ?, L'Harmattan, Paris.
- Direction du tourisme, 2009. Le champs du tourisme, http://www.tourisme.gouv.fr/fr/z2/stat/definitions/definitions_tourisme.jsp.
- Dreyfus-Signoles C., 2002. L'espace touristique. Bréal, 255 p.
- Dubois G., 2005. Indicators for an environmental assessment of tourism at national level. *Current Issues in Tourism*, vol. 8, n° 2-3.
- Dubois G. et Rechatin C., 2000. Tourisme, Environnement, Territoires : les indicateurs. IFEN, 260 p.
- Fiori A., Melot R. et Paoli J.C., 2008. L'aménagement du territoire à l'épreuve de la décentralisation : conflits et concertation en Corse et Sardaigne *revue Pôle sud*, n° 28, pp. 145-166.
- Flament E., 1975. Quelques remarques sur l'espace touristique. *Noroiis*, n° 88, pp. 609-621.
- Gartner W.C., 1993. Image Formation Process. in Uysal M. et Fesenmaier D.R. (eds.), *Communication and Channel Systems in Tourism Marketing*. The Haworth Press, London and New York, pp. 191-215.
- Girardin F., Calabrese F., Dal Fiorre F., Biderman A., Ratti C. et Blat J., 2008. Révéler la présence et les mouvements des touristes grâce au contenu généré par les utilisateurs. Forum International des Statistiques du Tourisme, Paris, http://www.forumstat.tourisme.gouv.fr/ftp/fr_S4_2_girardin.pdf.
- Gritti J., 1967. Les contenus culturels du Guide Bleu : monuments et sites "à voir". *Communications*, n° 10, pp. 51-64.
- Guilcher G., 2000. Naissance et développement du guide de voyage imprimé : du guide unique à la série, une stratégie de conquête des lecteurs ? in Chabaud G. et al. (eds.), *Les guides imprimés du XVIe au XXe siècle. Villes, paysages, voyages*. Belin.

- Guilcher G., 2004. Entre continuité et innovation, les guides touristiques en langue française de 1900 à 1960. *in* Breuillot M. et Beaufils T. (eds.), *Le Guide de voyage, au fil du Rhin et ailleurs*, Strasbourg, Université Marc Bloch Strasbourg II.
- Hunt J.D., 1975. Image as a Factor in Tourism Development. *Journal of Travel Research*, vol. 13, n° 3, pp. 1-7.
- INSEE, 2002. Note technique de l'INSEE sur la méthode de calcul de l'estimation de l'emploi salarié. 3 p., http://www.chiffres-tourisme-paca.fr/upload/Note_methodologique_emploi_touristique_INSEE.pdf.
- Jenkins O.H., 1999. Understanding and measuring tourist destination images. *International Journal of Tourism Research*, vol. 1, n° 1, International Journal of Tourism Research, pp. 1-15.
- Jenny J., 1996. Analyses de contenu et de discours dans la recherche sociologique française : pratiques micro-informatiques actuelles et potentielles. *in* Mangabeira W. (ed.), *Qualitative Sociology and Computer Programs : Advent and Diffusion of CAQDAS*. Current Sociology vol. 44, pp. 279-290.
- Kerbrat-Oricchioni C., 2005. Suivez le guide ! Les modalités de l'invitation au voyage dans les guides touristiques : l'exemple de l'« île d'Aphrodite ». *in* Baider F., Burger M. et Goutsos D. (eds.), *La communication touristique. Approches discursives de l'identité et de l'altérité*. L'Harmattan, pp. 133-150.
- Kérouanton M.-H., 1998. Les communes touristiques bretonnes : la plupart en bord de mer, quelques sites à l'intérieur des terres. *Octant*, n° 73, pp. 9-12.
- Kirat T., Torre A., Aviles A., Caron A., Lefranc C., Galman M., Melot R., Rialland C. et Salazar M.I., 2005. Modalités d'émergence et procédures de résolution des conflits d'usage autour de l'espace et des ressources naturelles. Analyse dans les espaces ruraux. *Rapport de recherche du Programme "Territoires, Environnement et nouveaux modes de gestion : la "gouvernance" en question"*. INRA, 250 p.
- Knafou R., 1997. Atlas de France. Tourisme et loisirs. *Atlas de France*, La documentation française, 126 p.
- Koshar R., 1998. "What Ought to Be Seen": Tourists' Guidebooks and National Identities in Modern Germany and Europe. *Journal of Contemporary History*, vol. 33, n° 3, pp. 323-340.
- KPMG, 2008. Suivi barométrique de l'image touristique de la destination Corse. Agence du Tourisme de Corse, http://www.visit-corsica.com/espace_pro/pdf/observatoire/annexes2008/barometre_2008.pdf.
- Ladwein R., 2002. Les modalités de l'appropriation de l'expérience de consommation : le cas du tourisme urbain. Centre Lillois d'Analyse et de Recherche sur l'Evolution des Entreprises, Lille.
- Leboucher S., 2007. Qui sont les champions... des guides de voyage: Le Journal du Net, <http://www.journaldunet.com/management/0707/match-guides-voyage/index.shtml>.
- Lefort I., 2007. Jean-Michel Dewailly, Tourisme et géographie, entre pérégrinité et chaos ? *Geocarrefour*, vol. 82, n° 4.
- Lepors A., 1999. Rapport à la ministre de l'Emploi et de la Solidarité et à la secrétaire d'Etat au Tourisme de Propositions pour l'amélioration de la situation sociale et professionnelle des travailleurs saisonniers du tourisme. 154 p., <http://www.tourisme.gouv.fr/fr/navd/mediatheque/publication/economie/att00006306/rapcomplet.pdf>.
- Lesbegueries J., 2007. Plate-forme pour l'indexation spatiale multi-niveaux d'un corpus territorialisé. Thèse en Informatique, Université de Pau et des Pays de l'Adour. Laboratoire d'Informatique, <http://tel.archives-ouvertes.fr/docs/00/25/85/34/PDF/Memoire.pdf>, 192 p.
- Livre Hebdo, 2007. Tourisme et voyage, Livre Hebdo.
- MacCannell D., 2001. Tourist agency. *Tourist Studies*, vol. 1, n° 1, pp. 23-37.
- MacGregor A., 2000. Dynamic texts and tourist gaze: Death, bones and buffalo. *Annals of Tourism Research*, vol. 27, n° 1, pp. 27-50.
- Mao P., 2003. Les lieux de pratiques sportives de nature dans les espaces ruraux et montagnards, Contribution à l'analyse de l'espace géographique des sports, Université Joseph-Fourier - Grenoble I. PACTE, <http://tel.archives-ouvertes.fr/tel-00280034/fr/>, 693 p.
- Martinetti J., 2007. Les tourments du tourisme sur l'île de Beauté. *Hérodote, revue de géographie et de géopolitique, numéro spécial Géopolitique du tourisme*, n° 127 -
- Melot R., Cadoret A., Jeanneaux P., Kirat T., Pham H.V. et Sabau C., 2008. Conflits environnementaux et gestion des espaces. Modalités de recours au tribunal, pratiques administratives et logiques d'acteurs entre environnement et aménagement. *Rapport de recherche*, INRA, GIP Mission de Recherche Droit et Justice, 133 p.
- Mercille J., 2005. Media effects on image: The Case of Tibet. *Annals of Tourism Research*, vol. 32, n° 4, pp. 1039-1055.
- Milner L.M., Collins J.M., Tachibana R. et Hiser R.F., 2000. The Japanese Vacation Visitor to Alaska: A Preliminary Examination of Peak and Off Season Traveler Demographics, Information Source Utilization, Trip Planning, and Customer Satisfaction. *Journal of Travel and Tourism Marketing* vol. 9, n° 1-2, pp. 43-56.

- Miossec J.M., 1977. L'image touristique comme introduction à la géographie du tourisme. *Annales de géographie*, n° 473, pp. 55-70.
- Moisy L., 2001. L'espace de la ville ludique et touristique : approche à travers les pratiques spatiales des visiteurs (hébergements, itinéraires) *Géocarrefour*, vol. 76, n° 2, pp. 107-113.
- Monereau M., 2008. Management des organisations touristiques. Bréal, 254 p.
- Perkins H.C. et Thorns D.C., 2001. Gazing or Performing?: Reflections on Urry's Tourist Gaze in the Context of Contemporary Experience in the Antipodes. *International Sociology*, vol. 16, n° 2, International Sociology, pp. 185-204.
- Petit D., 2007. Le guide de voyages, Caractère, le site des professionnels de l'imprimé, pp. 68-72.
- Prod'homme S., 2002. Voyager au Canada, la mise en route par les guides. *Alinéa*, vol. 1, n° 13, pp. 124-140.
- Quinlan S., 2005. 'Never short of a smile': A Content Analysis of Travel Guidebooks. *Recreation and Leisure Studies - Tourism Policy and Planning*, Waterloo, 271 p.
- Reinert M., 2003. Le rôle de la répétition dans la représentation du sens et son approche statistique par la méthode ALCESTE. *Semiotica*, vol. 147, n° 1-4, pp. 389-420.
- Richards G., 2002. Tourism attraction systems: Exploring Cultural Behavior. *Annals of Tourism Research*, vol. 29, n° 4, pp. 1048-1064.
- Roberts C.W., 1997. Text analysis for the social sciences. Methods for drawing statistical inferences from texts and transcripts. Lawrence Erlbaum Assoc Inc, 316 p.
- Santos C.A., 2004. Framing Portugal - Representational dynamics. *Annals of Tourism Research*, vol. 31, n° 1, pp. 122-138.
- Saunier P.Y., 1993. Le guide touristique, un outil pour une possible histoire de l'espace: autour des guides de Lyon 1800-1914. *Géographie et Cultures*, n° 13, pp. 35-54.
- Schockaert S., de Cock M. et Kerre E.E., 2008. Location Approximation for Local Search Services using Natural Language Hints. *International Journal of Geographical Information Science*, vol. 22, n° 3, pp. 315-336.
- Thévenot L., 1985. Les investissements de forme. Cahiers du Centre d'Etude de l'Emploi, PUF, pp. 72.
- Toulier B., 2000. L'influence des guides touristiques dans la représentation et la construction de l'espace balnéaire (1850-1950). in Chabaud G. et al. (eds.), *Les guides imprimés du XVIe au XXe siècle. Villes, paysages, voyages*.
- Tritz C., 2001. Tourisme et espace : l'arc méditerranéen. Une approche de la médiatisation. Thèse de Géographie, Université de Nice-Sophia Antipolis. UFR Espace et Culture, Nice, 265 p.
- Tritz C., 2003. La Côte d'Azur de 1952 à nos jours, la mutation d'un espace touristique et de son image. *Pays Vésubien*, n° 4, pp. 75-95.
- Tritz C., 2006. La valorisation touristique d'un territoire: entre acteurs et territoire: proposition de méthodes et applications. Colloque international "Développement territorial et tourisme durable, une approche par la valorisation touristique", Bogota, ULB, Universidad Externado de Columbia et Universidad del Atlantico.
- Urbain J.D., 1998. L'idiot du voyage, histoires de touristes. Payot et Rivages, 430 p.
- Urry J., 1990. The Tourist Gaze. Leisure and Travel in Contemporary Societies. Sage, London.
- Urry J., 1995. The tourist gaze and the environment. in Urry J. (ed.), *Consuming places*. Routledge
- Violier P., 2007. Tourisme et médias : regards d'un géographe. *Le Temps des Médias*, vol. 1, n° 8, pp. 159-170.
- Young M., 1999. The social construction of tourist places. *Australian Geographer*, vol. 30, n° 3, pp. 373-389.
- Zillinger M., 2006. The Importance of Guidebooks for the Choice of Tourist Sites: A Study of German Tourists in Sweden. *Scandinavian Journal of Hospitality and Tourism*, vol. 6, pp. 229-247.

Tableau 4 : bilan des données classiques sur le tourisme

Domaine	Indicateur	Information utilisée	Mode de constitution de la donnée	Unité spatiale élémentaire	Périodicité	Commentaires et limites	Références	
Offre d'hébergement	Nombre d'unités d'accueil	Inventaire communal 1998	Enquête auprès des municipalités	commune	Décennie	Faible périodicité, déclaratif	(Direction du tourisme, 2009)	
		Enquête annuelle INSEE	déclarations administratives des établissements classés et enquêtes des fédérations professionnelles depuis 2003 sur le parc non classé	commune	Année			
	Nombre de lits	Idem	idem	commune	Année			
	Taux de fonction touristique	Idem + RP9	Nb de lits/100 habitants	commune	Année + décennie			
	Densité touristique :	idem	Nombre de lits/ surface communale	commune	Idem			
	Indicateur de passage touristique	Idem	nombre de lits touristiques marchands / lits touristiques non marchands	commune	Année		(Kérouanton, 1998)	
Offre de services	Part de l'emploi saisonnier	INSEE-Sirene	Utilise les déclarations mensuelle UNEDIC10: nombre d'emplois salariés en été / nombre d'emplois en mars	commune	annuel	tourisme pas seul emploi saisonnier. estimations approximatives. petites entreprises et secteurs non couverts	(Kérouanton, 1998)8	
	Part de l'emploi touristique	idem	Idem + Méthode "Montpellier" : classement des familles d'activités selon l'intensité de leur lien au tourisme	commune	annuel	Méthodologie peu stabilisée, fourchette d'emploi large, inadaptée pour les grandes agglomérations, classement non universel des activités	(INSEE, 2002 ; Lepors, 1999)	
Fréquentation touristique effective	Nuitées d'hôtel	Enquête de fréquentation	Enquêtes mensuelles sur l'ensemble de l'année pilotée par les Directions régionales INSEE et partenaires locaux	?	mensuel	Exploitation conjoncturelle		
	Nuitées de camping	idem	Idem sur la saison estivale uniquement	?	mensuel	idem		
	Analyse des flux	Enquête aux frontières	Avant 1996, enquête sur séjours et dépenses des visiteurs étrangers aux lieux de sorties du territoire français, après 2004 réévaluation par modélisation	variable	variable		Instabilité due aux modifications des règles frontalières	(Direction du tourisme, 2009)
		Enquête SDT11	Sondage Sofres : déplacement et caractéristique des séjours des français d'un panel renouvelé au tiers/an (METASCOPE) depuis 1990 avec amendements en 1999 et 2004	commune ou département	Mensuel		autres déplacements que touristiques, sous-représentation de certaines destinations; difficulté d'identification des lieux, difficulté d'analyses diachroniques	(Ceron <i>et al.</i> , 2008)
Taxe de séjour	Ministère de l'Intérieur	Montant des taxes de séjour, assises sur les nuitées, perçues par les communes	commune	annuel	Choix communal, liée au touriste ou, depuis 1989, au logeur			
Classements administratifs	"Commune touristique"	Idem	Sur demande des municipalités, critère statistique sur le Taux de fonction touristique et présence d'un Office du tourisme	commune	5 ans	Lié au versement de la Dotation supplémentaire, réformes fréquentes	Loi 2006/437 (2006) , arrêté et décret dérivés	
	"Station classée"	idem	Présence de curiosités naturelles, pittoresques, historiques ou artistiques ou d'avantages du point de géographique, hydrominéralogique, climatique ou altitudinale	commune	5 ans	Cf loi de 1907 sur les casinos		
Qualification de l'attractivité touristique	"principales raisons d'attraction touristique"	Inventaire communal 1998	enquête déclarative auprès des municipalités : la municipalité classe les 3 premières raisons à partir d'une liste de 22 items	commune	décennie	Estimation des variables touristiques fragiles	(CERTU, 2007)	
	"Intensité touristique"	idem	Idem : appréciation de l'importance des touristes par rapport à la population permanente tous les mois	commune	décennie	idem	idem	
Image et notoriété des lieux	Notoriété spontanée	Sondages de type marketing	Pourcentage des lieux cités spontanément (une région, un département, un site suivant la nature du commanditaire)	Variable	variable	Non exhaustivité spatiale	(Monereau, 2008)	
	Notoriété assistée	idem	Pourcentage des lieux cités parmi une liste	Variable	Variable	idem	idem	

⁹ Recensement de la Population

¹⁰ Déclaration Annuelle des Données Sociales (DADS)

¹¹ "Suivi de la Demande Touristique des Français"