

> PLAGES <



ACTUALISATION DE LA FRÉQUENTATION DES PLAGES DE NOUVELLE-AQUITAINE



RAPPORT PUBLIÉ
EN NOVEMBRE 2023

GIP LITTORAL

2030





Table des matières

Introduction	4
Partie 1 Valorisation de la donnée existante pour une première approche de la fréquentation au niveau régional.....	6
1.1 Analyse des données disponibles	7
1.1.1 Analyse de l'étude de fréquentation menée en 2017	7
1.1.2 Récapitulatif de la robustesse de l'étude de 2017.....	11
1.1.3 Autres données analysées	14
1.2 Estimation de la fréquentation des plages en 2022 sur la Nouvelle-Aquitaine.....	17
1.2.1 Proposition d'une méthode et d'hypothèses alternatives et complémentaires	17
1.2.2 Récapitulatif des robustesses	21
1.2.3 Mise en œuvre du modèle Nouvelle-Aquitaine 2022	22
1.2.4 Résultats : Étendue des données de 2017 au périmètre de Nouvelle-Aquitaine.....	32
1.2.5 Résultats : Étendue des données de fréquentation à 2022 en Nouvelle-Aquitaine...	35
Partie 2 Proposition de relance d'une campagne de fréquentation	38
2.1 Proposition méthodologique	38
2.1.1 Stratégie de déploiement.....	39
2.1.2 Chronologie de déploiement	40
2.2 Estimation technique et financière.....	41
2.2.1 Support de Chiffrage.....	41
2.2.2 Pilotage et suivi du projet.....	42
2.2.3 Outil de simulation d'une campagne de fréquentation des plages.....	43
Conclusion	44
Annexes.....	46
Annexe 1 : Variables de caractérisation des plages proposées	46
Annexe 2 : Analyse du « silhouette score » pour 3 clusters	48
Annexe 3 : Liste des plages par clusters	49
Annexe 4 : Formules de calcul.....	54

Introduction

Le présent rapport d'étude s'inscrit dans le cadre de l'animation par le GIP Littoral du schéma régional plans plages. Validé en 2010, le schéma Plan Plage est l'un des premiers chantiers lancés par le Conseil d'Administration du GIP. La finalité du schéma est la coordination de tous les projets d'aménagement des plages de Nouvelle-Aquitaine en appliquant des principes communs et tout en s'adaptant aux situations locales. Ainsi, à côté de sa mission d'accompagnement des porteurs de projets de démarches en matière d'aménagement durable des plages, le GIP Littoral anime un certain nombre d'actions collectives (postes de secours mobile en bois, panneaux réglementaires baignade, production de notes techniques...). Le travail conduit en 2022 sur l'opportunité de relance d'une campagne de fréquentation des plages à l'échelle de la Nouvelle-Aquitaine s'inscrit dans ce cadre.

[Une première campagne de fréquentation régionale des plages a été réalisée en 2017](#) par le GIP Littoral, accompagné de la société **ANTS**¹, avec pour objectif d'estimer le nombre de visiteurs sur les 128 plages surveillées du littoral des trois départements d'ex-Aquitaine.

Cette campagne a permis au partenariat et aux acteurs du littoral de bénéficier d'une première donnée de référence sur le nombre de visites sur les plages du littoral régional (10 millions de visites² sur les plages de juillet à août), dont l'utilité s'est manifestée à deux échelles :

- Au niveau régional et départemental, pour appuyer les politiques publiques existantes en matière de réaménagement durable des plages ;
- Au niveau local, pour chaque site étudié, via la remise d'une fiche sur la fréquentation de la plage pour répondre aux besoins exprimés en matière d'observation et de gestion des sites.

Cette donnée a été utilisée par un panel diversifié de partenaires : acteurs du tourisme, gestionnaires des plages, acteurs de la santé, etc.

[L'opportunité de relancer une campagne de fréquentation](#) visant à actualiser la donnée sur la fréquentation des plages a été validée en Groupe technique Plages en 2021 et inscrite au programme de travail du GIP en 2022. Les besoins exprimés étaient les suivants :

- Au niveau régional : disposer d'une donnée actualisée et étendue à l'échelle du périmètre du littoral de la Nouvelle-Aquitaine ;
- Au niveau local : un renforcement des besoins d'observations.

De manière générale, le besoin de renforcer la méthodologie de caractérisation des plages proposée en 2017, et notamment en y intégrant la diversité des plages de Charente-Maritime a été exprimé. Le groupe technique a proposé de se fixer de nouveaux objectifs si une campagne de fréquentation devait être relancée :

- Renforcer utilisation de la donnée par le plus grand nombre (exemple : analyse d'évènement sanitaire ...) ;
- Pérenniser un dispositif pour mesurer l'évolution sur le temps long (renouvellement tous les 5 ans par exemple...).

La question préalable à l'opportunité de relancer une campagne de fréquentation à l'échelon du nouveau périmètre régional (avec mise en place de dispositifs de comptage) a été celle de pouvoir, dans un premier temps, **étudier la possibilité de capitaliser sur la donnée déjà existante**, afin de **d'inscrire ce possible renouvellement de campagne de comptage dans le temps long**.

¹ Société spécialisée en conception de modèles prédictifs, traitement de très grands flux de données et développement de capteurs intelligents.

² Pour précision, il s'agit bien ici du nombre de visites et non du nombre de visiteurs (si un usager est allé plusieurs fois à la plage dans la saison, toutes ses visites ont été comptabilisées)

La mission a été confiée en novembre 2022 à la société **Hupi**, spécialisée dans l'application de l'Intelligence Artificielle au service des besoins opérationnels, et plus récemment sur les questions d'affluence³. Elle s'est conduite en 2 phases :

- Une 1^{ère} phase de valorisation de la donnée existante pour une première approche de la fréquentation au niveau régional

L'objectif de la 1^{ère} phase était d'évaluer et de lever les hypothèses permettant de construire l'estimation de la fréquentation au niveau régional sur la base des données existantes/accessibles, afin de "fiabiliser les chiffres" et de fournir des indicateurs de tendances relatif à la saison touristique. Pour ce faire, un travail préalable d'identification et d'évaluation du potentiel d'utilisation des données existantes a été mené, avec pour objectif d'évaluer l'impact de chacune de ces données sur le niveau de fréquentation. Les données fournies au démarrage de la mission n'étant pas exhaustives, il était attendu un soutien du prestataire à l'identification d'autres sources de données ayant un pouvoir prédictif, et qu'il pourrait être utile de mobiliser. Le présent rapport identifie et analyse ces données dans la perspective de valider ou invalider leur fiabilité, et voir dans quelle mesure elles influent sur la fréquentation des plages.

Ensuite, le rapport propose une première approche actualisée de la fréquentation des plages au niveau régional en 2022, par l'extrapolation des résultats de 2017 à l'année 2022, à l'aide d'indicateurs sociaux et macro-économiques récupérés via l'INSEE.

L'objectif était d'évaluer la qualité des prévisions, basées sur des données variables sans mesures sur terrain.

- Une 2^{ème} phase de proposition d'organisation de la campagne de fréquentation 2024

L'objectif de la 2^{ème} phase était de proposer les éléments sous-jacents à l'organisation de la campagne : sites à équiper, caractéristiques des dispositifs et solutions technologiques mises en œuvre, organisation générale de la campagne, coûts estimés, etc.

Le schéma ci-dessous synthétise le point de départ (en orange) et les travaux réalisés (en bleu) pour répondre aux objectifs.



Le présent document constitue le rapport final de l'étude sur l'actualisation 2022 de la fréquentation des plages de la région Nouvelle Aquitaine.

³ Référence projet : « Gestion des flux touristiques en zone de montagne » – ADT64

Partie 1 | Valorisation de la donnée existante pour une première approche de la fréquentation au niveau régional

Le travail a été conduit en plusieurs étapes :

- **Reprise du travail** d'analyse réalisé lors de l'étude de 2017, recensant et caractérisant les hypothèses faites lors de cette étude (tant d'un point de vue des outils techniques que des valeurs remontées). Ce travail récapitule la robustesse du travail mené en 2017 selon une méthodologie spécifique explicitée ci-après (Acquisition – Segmentation – Consolidation).
- **Etude et évaluation** des études réalisées en parallèle avec des méthodes différentes, notamment sur La Rochelle, ainsi que sur les comptages effectués sur les parkings des plages de Gironde. L'objectif est ici de capitaliser sur ces recherches et ces données en prenant en compte les remarques et observations concernant les différentes méthodologies mises en place.

Proposition et description d'une méthodologie et des hypothèses alternatives permettant d'affiner l'estimation de la fréquentation des plages, par la combinaison d'informations relevées sur le terrain et l'application d'outils mathématiques (Intelligence Artificielle) pour mettre en oeuvre une stratégie adaptée et ciblée permettant de couvrir l'ensemble des plages.

Pour répondre à ces objectifs, trois problématiques ont été identifiées (basées sur une méthodologie d'analyse de données en Intelligence Artificielle et statistiques avancée, développée par Hupi) :

1. **Acquisition des données** : acquisition des données externes (dites « open-source ») et des données de la campagne de fréquentation 2017.
2. **Segmentation** : caractérisation de l'ensemble des plages, ainsi que des résultats d'un regroupement mathématique des plages selon leurs caractéristiques communes (« clustering »).
3. **Consolidation** : mise en oeuvre des modèles statistiques de fréquentation des plages, couplée à une « triangulation des données » (processus qui permet de contrôler les résultats des données d'origine à l'aide de l'ajout de plusieurs sources de données différentes).

Enfin, les résultats ont été extrapolés de 2017 à 2022 sur l'ensemble des plages de Nouvelle Aquitaine, en prenant en compte les données fournies par le GIP Littoral ainsi que certaines données open source.

Définitions :

Il est important avant toute chose de définir ce qu'est une plage dans le contexte de ce projet. Une plage correspond ici, et dans la suite de ce document, à une plage surveillée. Les plages pour lesquelles il y a plusieurs zones de baignade sont comptabilisées comme une seule plage car le raisonnement se fait par poste de secours. Ainsi, 189 plages surveillées (plages océanes et plages lacustres) sont comptabilisées sur le périmètre régional de Nouvelle Aquitaine.

1.1 | Analyse des données disponibles

1.1.1 | Analyse de l'étude de fréquentation menée en 2017

Rappel de l'approche méthodologique utilisée en 2017

En 2017, le GIP Littoral, accompagné de la société ANTS, a réalisé une première étude de portée régionale pour estimer le nombre de visites sur les 128 plages surveillées de l'ex Aquitaine.

Après une 1^{ère} phase de test technique conduite en 2016 sur 5 plages, le dispositif a été étendu en 2017 à 18 plages. Le choix s'est porté sur des capteurs permettant de mesurer en temps réel le nombre de smartphones présents dans un rayon d'un kilomètre via un système de détection des ondes wifi. 220 000 mesures ont été réalisées.



Les 18 plages choisies ont été sélectionnées de la manière suivante : au départ le choix a été fait d'équiper 11 plages en Gironde⁴, 4 plages dans le département des Pyrénées-Atlantiques et 3 plages dans le département des Landes. Chaque plage équipée comportait des caractéristiques différentes. Sur la côte basque, on note pratiquement une plage par commune.

En parallèle de l'installation des capteurs, 49 survols de drones ont été effectués, couplés au déploiement de méthodes de « deep learning⁵ » permettant de compter automatiquement le nombre de personnes présentes sur chaque image. Les résultats ont été très bons lorsque la concentration des personnes n'a pas été très importante. En revanche, à partir d'une certaine densité, le taux d'erreur a dépassé les 10%.

Un 1^{er} ratio entre le maximum de personnes au pic de la journée identifié par les capteurs et le nombre de visites uniques a été calculé en 2016 (valeur : 2.55). Celui-ci a été requestionné en 2017, dans la mesure où les comptages ont mis en évidence que ce ratio n'était pas constant et variait selon la topologie de la plage ainsi que l'affluence. Les 5 plages étudiées en 2016 étant plus ou moins équivalentes, il était important de vérifier ce résultat en l'appliquant à différentes plages.

Les données étudiées ont permis d'avoir accès au pic de fréquentation maximum de la journée. Il aurait été possible d'obtenir le nombre de visiteurs directement avec les capteurs, puisque chaque téléphone a été identifié de manière unique. Cependant, comme il est interdit de conserver une donnée personnelle plusieurs heures, la seule solution était alors d'étudier les boucles de comptage ONF des aires de stationnement.

Le ratio de 0.64 a été obtenu (lorsque le nombre maximum de personnes de la journée est de 640, 1000 personnes sont venues à la plage ce jour-là).

⁴ En lien avec la mise en place de l'application test « Go Plages » en 2016 : au-delà de l'observation de la fréquentation, enjeu de tester la remontée d'information sur la fréquentation de certaines plages de Gironde plus ou moins fréquentées, auprès d'un panel d'utilisateurs ciblé.

⁵ Procédé d'apprentissage automatique s'appuyant sur des neurones artificiels.

Toutefois, ce ratio reste à utiliser avec prudence puisqu'il a été calculé à partir de seulement 4 parkings de sites de plages sur lesquels la part de visiteurs en voiture est très largement dominante (au-delà de 90%).

D'autres biais seraient également potentiellement à prendre en compte : le caractère de « plage nature » des 4 plages de Gironde dont les données ont été étudiées.

A savoir qu'une plage nature correspond à une plage complètement isolée de la ville, à laquelle on ne peut pas accéder facilement à pied ou à vélo.

En parallèle de l'étude des données issues des 18 capteurs, une caractérisation des plages du littoral aquitain a été réalisée afin de regrouper les plages et de pouvoir extrapoler les résultats obtenus sur l'ensemble des plages du littoral aquitain. Les caractérisations se sont basées sur plusieurs données de description :

- Données topologiques (étendue, capacité d'accueil, accessibilité, situation en zone urbaine ou naturelle)
- Variables qualitatives : lacustre, privée, naturaliste, urbaine ou nature, avec pression touristique et/ou métropolitaine
- Données socio-économiques : population en résidence principale de la commune, population en résidence secondaire de la commune, capacités d'accueil touristique de la commune, nombre de plages de la commune, nombre de commerces liés à l'activité touristique (< 1 km), nombre de parkings et campings et capacité d'accueil liée à un seul établissement (plage privée)

En 2017, les deux mesures clés définies ont ainsi été :

- Le pic au maximum de la journée (nombre de personnes maximum sur la plage)
- Le nombre de visites

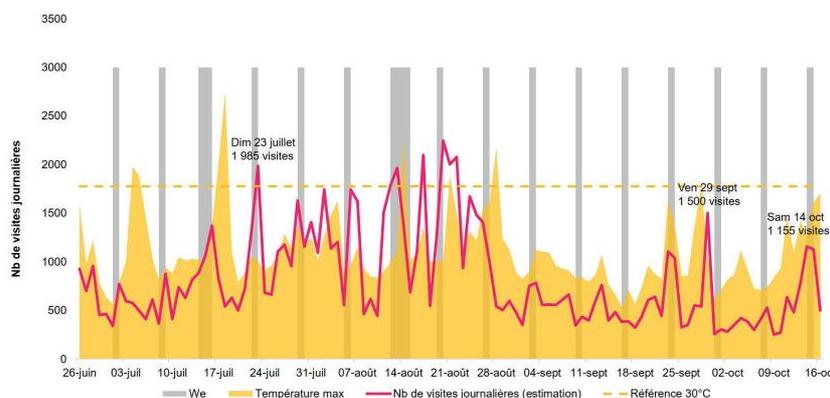
La campagne conduite en 2017 a mis en évidence que la température et l'affluence évoluent conjointement (elles sont corrélées). Le modèle est ainsi plus performant dans le cadre de la prise en compte des données météorologiques. La température influencerait positivement et le niveau de précipitation influencerait négativement sur l'affluence. Ces résultats sont rassurants puisqu'ils sont intuitifs.

Ces prédictions ont pu être interrogées de deux manières différentes : pour toute la journée ou pour une heure précise.

En parallèle, les mesures ONF provenant de capteurs « éco-compteur » comptant les entrées et sorties des véhicules ont également été étudiées. Une corrélation a bien été détectée entre les mesures ONF et les mesures des capteurs ANTS. Cependant, ce rapport est variable par plage et par jour.

Illustrations des résultats 2017 issus de la fiche de synthèse du GIP :

COURBE CROISEE FREQUENTATION / METEO

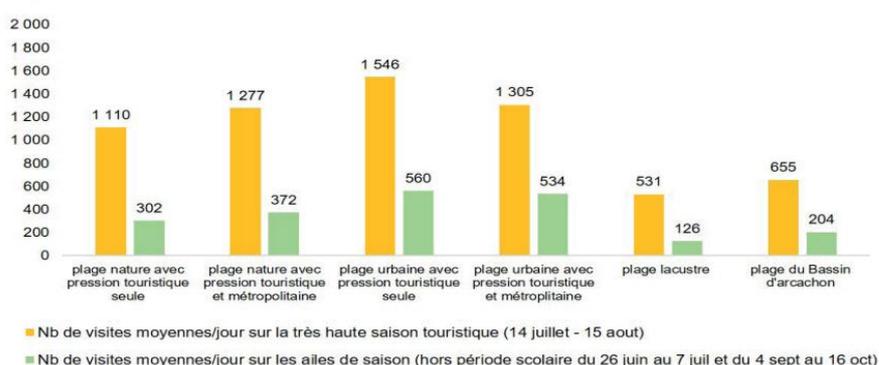


Extrait de la fiche plage de Bidart : courbe croisée fréquentation/météo saison estivale 2017

FREQUENTATION MOYENNE PAR DEPARTEMENT ET SAISON



FREQUENTATION MOYENNE PAR TYPOLOGIE DE SITE ET SAISON



Exemples de résultats obtenus en 2017 : fréquentation moyenne selon le département/la typologie de site et la saison (haute saison – ailes de saison – 3 jours les plus chargés)

SYNTHESE

	Juillet	Aout	Septembre
Maximum journaliers observés en Gironde	830 000	1 270 000	255 000
Maximum journaliers observés dans les Landes	1 030 000	1 515 000	350 000
Maximum journaliers observés dans les Pyrénées Atlantiques	645 000	890 000	330 000
Total de la fréquentation journalière maximum observée	2 505 000	3 675 000	935 000
Estimation du nombre de visites maximum	3 932 850	5 769 750	1 467 950
Estimation du nombre moyen de visites par jour	835	1 265	310

Synthèse des résultats obtenus en 2017

Analyse critique des hypothèses formulées durant l'étude de 2017

Regroupement des plages "similaires" pour l'extrapolation : ce regroupement a été fait sur les données topologiques et les données socio-économiques fournies par le GIP Littoral. Les plages ont ainsi été regroupées de la manière suivante :

- Plages du Bassin d'Arcachon
- Plage lacustre avec pression touristique seule
- Plage nature avec pression touristique et métropole
- Plage nature avec pression touristique seule
- Plage urbaine avec pression touristique et métropole
- Plage urbaine avec pression touristique seule
- Plage non surveillée
- Plage privée (naturiste ou centre de vacances)

Définition de la pression touristique et de la présence de métropole :

La pression touristique⁶ a été prise en compte pour l'ensemble des plages.

La proximité avec une métropole (ou agglomération) a été prise en compte pour les plages situées entre Lacanau et Biscarosse, sur toute la côte basque et sur le sud des Landes.

Nb : La notion de métropole abordée ici (et dans la suite du rapport) correspond plutôt à la présence d'une forte entité urbaine qui constitue un bassin émetteur de fréquentation (grande ville, agglomération, Métropole...) à proximité de la plage étudiée.

Restriction au niveau des données météorologiques :

Seule la tendance journalière a été prise en compte et non la tendance heure par heure, car celle-ci ajoutait un « bruit statistique⁷ ». Les données météorologiques étudiées étaient seulement la température et le niveau des précipitations (la couverture nuageuse par exemple n'était pas prise en compte car seuls la température et le niveau de précipitations ont été considérés comme étant les facteurs les plus influents pour la société ANTS).

Pour le littoral girondin, les données météorologiques prises en compte étaient celles de la station de Bordeaux. En conséquent, des variations ont été remarquées (pas automatiquement la même température sur la station de Bordeaux que sur le littoral girondin).

Seule la période estivale a été étudiée par le modèle (pas d'étude sur la période hivernale). En conséquent, lorsque le modèle a effectué sa phase d'apprentissage, il n'a pas pu parcourir toutes les gammes de températures qu'il aurait pu rencontrer (les conditions météorologiques en 2017 étaient alors particulières, avec des pics de chaleur réguliers mais des week-ends en haute saison et ailes de saison plutôt décevants).

Le calendrier des festivités, des jours fériés et des ponts n'a pas été pris en compte, le modèle a donc eu tendance à sous-estimer les prédictions sur ces jours-là.

Le ratio de 0.64 entre le maximum de personnes au pic de la journée et le nombre de visites uniques ayant été calculé à partir des données de stationnement de seulement 4 plages, il est à prendre en compte avec prudence, puisqu'il n'a pas été validé mathématiquement et les 4 plages équipées sont 4 plages natures de Gironde. La société ANTS avait émis l'hypothèse que ce même ratio ne serait pas applicable aux plages du Pays Basque par exemple.

D'autres données (comme les marées par exemple) n'ont pas été intégrées au modèle en 2017. La suite du rapport étudie d'ailleurs la possibilité d'intégrer de nouvelles données au modèle.

Caractérisation de la robustesse des hypothèses de 2017 et priorisation

Cette partie reprend les hypothèses abordées précédemment, mais permet de les caractériser grâce au tableau ci-dessous.

Pour chaque hypothèse (chaque ligne) nous retrouvons 3 colonnes permettant de les caractériser :

- La première colonne correspond à un **niveau d'impact** (faible/moyen/fort) sur l'objectif (ici, l'estimation de fréquentation des plages)
- La seconde colonne correspond à la **source de l'hypothèse** : hypothèse extraite de données open source / de connaissance métier / d'étude.
- La dernière colonne correspond à la **sensibilité**. Ici on associe un niveau de sensibilité de l'hypothèse qui permet de définir s'il est nécessaire d'affiner l'hypothèse, soit via une étude, soit via un apport de données.

⁶ Le taux de fonction touristique (également appelé indicateur d'intensité touristique) est le rapport entre la capacité en hébergements touristiques des communes (nombre de lits touristiques) et leur population résidente à l'année. Il s'agit d'un indicateur de pression touristique permettant de quantifier la multiplication théorique de la population en période d'afflux touristique.

⁷ Données imprévisibles et sans rapport avec les variables étudiées. Le bruit statistique est le résultat d'erreurs de mesure, de facteurs externes aléatoires et/ou de processus internes aléatoires. <https://commentouvrir.com/definitions/bruit-statistique/>

Hypothèses	Impact sur l'objectif (priorisation des hypothèses)	Source de l'hypothèse (nécessité d'études ou validation)	Sensibilité (nécessité d'affiner l'hyp via étude ou apport de données)
Classification des plages (données topologiques et données socio-économiques)	Fort	Métier	Moyenne
Choix pression touristique	Moyen	Métier	Faible
Choix présence de métropole	Moyen	Métier	Forte
Données météorologiques à la journée	Fort	Open source	Faible
Sélectionner uniquement la température et le niveau de précipitations	Fort	Open source	Forte
Utilisation des données météorologiques de la station de Bordeaux pour les plages du littoral girondin	Moyen	Open source	Forte
Étudier uniquement la saison estivale	Faible	Métier	Moyenne
Calendrier des festivités et des jours fériés et ponts non pris en compte	Moyen	Open source	Forte
Considérer que le ratio entre le nombre de personnes maximum et le nombre de visiteurs est constant	Fort	Etude 2016	Forte
Considérer le ratio de 0.64 comme applicable à toutes les plages du littoral	Fort	Etude compteurs parkings	Forte

1.1.2 | Récapitulatif de la robustesse de l'étude de 2017

Acquisition des données

Les données externes acquises lors de l'étude de 2017 sont les suivantes :

- Les données météorologiques : température maximum.
- Les données calendaires avec notamment le jour de la semaine, le jour de l'année, le mois et l'heure.
- Les données topologiques : plage lacustre, privée, naturiste, urbaine ou nature, avec pression touristique seule et/ou métropolitaine.
- Les données socio-économiques : population en résidence principale de la commune, population en résidence secondaire de la commune, capacités d'accueil touristique, nombre

de plages de la commune, nombre de commerces liés à l'activité touristique, nombre de parkings et campings.

La campagne de fréquentation de 2017 a permis de mesurer approximativement la fréquentation des plages, mais comme abordé précédemment, le système utilisé comportait certaines limites. Le protocole de mesure se basait sur le comptage (via sa clé wifi) du nombre de paquets provenant d'appareils différents avec un périmètre de détection non défini.

Tout d'abord, le périmètre de détection ne pouvait pas être défini puisque selon les composants des appareils, les paquets sont envoyés plus ou moins loin. De ce fait, le nombre d'appareils sur la plage peut être plus important que le nombre de paquets reçu par les capteurs.

D'autre part, le protocole mis en place ne différenciait pas les types d'appareils émettant des paquets wifi. De même pour les familles se rendant à la plage, ce protocole n'a pas pu détecter les enfants (sachant qu'en moyenne en France les couples ont 1,87 enfants).

De plus, toutes les personnes se rendant sur les plages ne sont pas forcément équipées d'appareil émetteur, ce qui a pu donner lieu à un comptage inférieur au nombre de personnes réellement sur la plage. Enfin, les utilisateurs de smartphones n'ont pas forcément le réseau wifi activé lorsqu'ils sont sur la plage. Ce type de réseau est principalement utilisé à domicile ou dans des hôtels ou campings, alors que sur la plage, c'est le réseau 4G qui est généralement privilégié.

Aussi, la notion de visiteur unique était difficile à définir.

Segmentation

La caractérisation des plages effectuée pour la campagne de 2017 se basait uniquement sur 4 variables principales avec, tout d'abord, la distinction des plages du Bassin d'Arcachon, et ensuite, une caractérisation plus fine pour les plages hors Bassin d'Arcachon qui prenait en compte la pression touristique, la pression métropole ainsi que la nature de la plage (lacustre, nature, urbaine). Ainsi, cette caractérisation ne semble pas très robuste, puisqu'elle part d'une analyse logique, mais ne prend en compte que peu de variables et ne prend pas en compte les corrélations entre les différentes plages.

Cette segmentation peut être conservée, mais il faudra dans les études mettre en place des clusters qui seront abordés par la suite.

Clustering (regroupement des plages par caractéristiques) : Il est aujourd'hui considéré que le clustering effectué lors de l'étude de 2017, se basant sur la première étape de caractérisation des plages, pourrait être plus robuste si d'autres variables étaient intégrées à la caractérisation.

De plus, ce clustering n'a, par la suite, pas été utilisé dans le modèle de prédiction de l'affluence. Ce sont les données brutes qui ont été considérées dans la matrice d'entrée pour le modèle statistique.

Un des objectifs dans le cadre de la présente étude a été de renforcer cette étape de clustering, afin de pouvoir l'intégrer au modèle d'estimation de la fréquentation 2022.

Consolidation

Modèle statistique : La matrice d'entrée du modèle de 2017 contenait un nombre important de variables puisqu'elle recensait les données calendaires, les données météorologiques ainsi que les données socio-économiques associées à chaque plage.

Le modèle statistique de « gradient boosting⁸ » intégré en 2017 a montré de bons résultats sur les indicateurs clés (R^2 , MAE, RMSE)⁹. Cependant, un doute subsiste sur la construction de ce modèle

⁸ Modèle d'apprentissage automatique utilisé pour créer des modèles prédictifs.

⁹ R^2 : Le coefficient de détermination R^2 est une mesure, sans unité, de la corrélation des caractéristiques par rapport à la cible. Il peut être exprimé comme suit : à quel pourcentage plus proche de la perfection se trouvent les prédictions du modèle par rapport à l'utilisation de la valeur moyenne de la cible ? MAE : L'erreur absolue moyenne (Mean Absolute Error ou MAE) est la moyenne de toutes les erreurs de prédiction absolues, où l'erreur de prédiction est la différence entre la valeur réelle et la valeur prédite. RMSE : La racine de l'erreur quadratique moyenne (Root Mean Squared Error) peut être interprétée comme la différence +/- moyenne prévue entre une valeur prédite et la valeur réelle. Il s'agit de l'écart type des résidus (la différence entre la valeur observée et la valeur prédite d'une caractéristique).

statistique et notamment sur la séparation des données en 2 étapes pour le modèle, que sont la phase d'apprentissage et la phase de test. Ceci a pu produire un surapprentissage du modèle, qui a obtenu d'une part de très bons résultats sur les données utilisées pour son entraînement, mais d'autre part de mauvais résultats sur de nouvelles données, ce qui expliquerait les écarts obtenus une fois le modèle testé en temps réel.

Triangulation des données : Sur l'étude de 2017 il n'y a pas eu de triangulation des données, notamment sur des variables comme la pression métropole par exemple. **La triangulation est l'utilisation ou la recherche de plusieurs sources ou métriques pour approfondir la compréhension d'un phénomène.** Dans le cas de la pression métropole, cela consisterait à prendre en compte diverses variables permettant de la caractériser, et ainsi valider son pouvoir prédictif et améliorer la qualité des résultats.

Conclusion

Etapes	Robustesse
Acquisition des données : 1. Données externes 2. Campagne de fréquentation	Forte Forte
Segmentation : 1. Caractérisation des plages 2. Clustering	Faible Très faible
Consolidation : 1. Modèle statistique 2. Triangulation des données	Faible Très faible

Points à retenir :

- Complexité d'évaluation et de validation des données pour obtenir des informations sur les volumes de fréquentation (doutes importants).
- La technologie utilisée limitante, mais des indicateurs de redressement pertinents pour une plage donnée.
- Difficulté d'extrapolation pour toutes les plages. Une première caractérisation des plages faite en 2017, mais à approfondir en ajoutant d'autres variables.
- Un doute sur la mise en place du modèle fait par ANTS en 2017 et notamment la séparation des données pour la phase d'apprentissage et la phase de test.
- De nouvelles approches à considérer, telles que la mise en place d'un clustering et la triangulation des données.

Ces 3 indicateurs sont appelés des « métriques de régression », elles sont utilisées dans la construction de modèles d'apprentissage pour optimiser les modèles prédiction d'une variable quantitative, évaluer leurs performances finales et les comparer entre eux.

Source : https://help.qlik.com/fr-FR/cloud-services/Subsystems/Hub/Content/Sense_Hub/AutoML/scoring-regression.htm#anchor-2

1.1.3 | Autres données analysées

La société HUPI a également analysé d'autres sources de données disponibles qui n'ont pas été retenues dans la suite des travaux : ces données n'étant pas disponibles pour l'ensemble des plages, il a été difficile de les intégrer. Une donnée homogène est requise pour permettre une application au modèle et ainsi le rendre fiable.

Comptages réalisés sur les parkings de Gironde

Le dispositif de comptage en temps réel des parkings de plages déployé en 2021, avant d'être reconduit et consolidé en 2022 et 2023 a été pris en considération.

Ce dispositif constitue le premier axe opérationnel qui découle de la stratégie d'accueil sur les plages de Gironde proches de la Métropole bordelaise¹⁰, et permet une remontée d'information en temps réel et en prédictif sur le remplissage des parkings de cinq plages sur les communes du Porge, de Lège-Cap-Ferret, de Lacanau et, depuis 2022, de Carcans-Océan qui a rejoint le dispositif.

Ainsi, de la donnée sur la fréquentation des parkings de plage en Gironde en saison estivale est disponible depuis 2021.

Point à retenir :

- Fiabilité des solutions techniques et disponibilités des données. Cependant, des données qui ne concernent que 5 plages, donc non-intégrables au modèle. Ces données pourraient par ailleurs être utilisés pour vérifier les résultats du modèle.

Études effectuées sur La Rochelle

Un projet a été mené par l'Université de La Rochelle sur la fréquentation des plages de 2008 à 2014. Cette étude est basée sur des photos aériennes et des comptages qui ont été effectués sur les plages de l'Île de Ré, l'Île d'Oléron, les principales plages de La Rochelle, à Fouras et également à l'Île d'Aix. Les données collectées sont disponibles sur le serveur de l'UMR LIENSs.

D'autres projets ont été menés et ont nécessité la connaissance de la fréquentation des plages :

- 2014 : réflexion sur le remplissage des plages (à partir de prises de vue, estimation journalière et webcam)
- Entre 2019 et 2020 : étude de l'impact de la crème solaire sur l'environnement
- 2021 : étude de l'effet Covid sur la fréquentation des plages

Dans ces études, 3 types de plages ont été pris en compte :

- Plage urbaine (front de mer rattaché à un centre-ville)
- Plage de station (front de mer non rattaché à un centre-ville)
- Plage naturelle (sans arrière plage directe ou fortement construit)

Il a été observé que les rythmes de fréquentation sont différents selon les types de plages. Au contraire, l'étude du GIP Littoral menée en 2017 ne montrait pas de rythme différent, juste un décalage d'une heure pour le pic d'affluence entre les plages lacustres et les plages océaniques.

¹⁰ <https://www.giplittoral.fr/ressources/strategie-daccueil-des-plages-de-gironde-proches-de-la-metropole>

Entre 2008 et 2014 des enquêtes de fréquentation ont également été menées en parallèle des images aériennes. Chaque année, une dizaine de plages étaient enquêtées.

Il est à noter qu'il n'y a pas eu d'évolution de la fréquentation, seulement des variations inter-annuelles et inter-mensuelles (fréquentations différentes en fonction des années et des mois).

Les plages de La Rochelle et de Chatelaillon ont été équipées de webcam avec mémoire. Sur ces plages-là, le comptage a été effectué à partir des images de webcam et d'un maillage de la plage mis en place grâce à des drapeaux (permettant de compter le nombre de personnes par maille).

La densité sur les plages a également été étudiée : aucune densité n'a été identifiée comme étant problématique (pas de densité inférieure à 4m² par plageur).

L'urbanisation a également été prise en compte via la densité de bâtiments dans un rayon en arrière de la plage. Ce critère a également servi à qualifier les différentes plages.

Pour cette étude, la notion de plage n'est pas la même que pour l'étude menée en 2017 par le GIP Littoral. Une plage ne correspond pas à un poste de secours puisque certaines plages non-surveillées accueillent de la fréquentation et de la baignade (par exemple à La Rochelle).

Point à retenir :

- Une étude alternative basée sur d'autres outils techniques avec des résultats satisfaisants mais nécessitant un travail humain important, donc non reproductible à une échelle plus importante. Une information plus fine pour mesurer les pics d'affluence et les flux plus précisément.



Visuel issu de l'Atlas de la fréquentation des plages de Charente-Maritime - ECOP



Plage d'Aytré – La Rochelle (17)

Synthèse de l'étude 2017 et points à retenir :

- Une première étude réalisée en 2017 de portée régionale pour estimer le nombre de visites sur les 128 plages surveillées de l'ex-Aquitaine :
- Une collecte de données de terrain via des capteurs détectant les ondes wifi permettant de mesurer en temps réel le nombre de smartphones présents dans un rayon d'un kilomètre, installés sur 18 plages représentatives des différentes typologies de plages à l'échelle de la région. Couplé à un redressement de la donnée par survols de drones et le déploiement de méthodes de « deep learning ».
- Une caractérisation des plages du littoral aquitain, basée sur plusieurs données de description (topologies, qualitatives, socio-économiques), afin de regrouper les plages et extrapoler les résultats obtenus sur l'ensemble des plages du littoral aquitain.
- L'ajout au modèle des données météorologiques (températures, précipitations), influant sur la fréquentation.
- Certaines données ayant une influence, non prises en compte dans le modèle de 2017 : marées, période hivernale, calendrier des festivités, des jours fériés et des ponts.
- Une évaluation réalisée par la société Hupi sur la robustesse globale de l'étude réalisée en 2017, au regard des données utilisées, de la méthode de regroupement des plages, et du modèle statistique d'estimation. Un certain nombre de limites trouvé au modèle de 2017.
- L'objectif pour Hupi dans le cadre de la suite de l'étude a été de proposer une réutilisation des données et du modèle de 2017, avec des améliorations, afin d'actualiser l'estimation de la fréquentation pour 2022, sur le périmètre de la Nouvelle-Aquitaine.

1.2 | Estimation de la fréquentation des plages en 2022 sur la Nouvelle-Aquitaine

1.2.1 | Proposition d'une méthode et d'hypothèses alternatives et complémentaires

La partie suivante propose et décrit une méthode et des hypothèses alternatives pour faire estimer la fréquentation, à la fois à l'échelle du nouveau périmètre régional, et à la saison 2022. Comme mentionné précédemment dans le rapport, cette proposition méthodologique est construite en reprenant les trois problématiques identifiées d'acquisition des données, de segmentation et de consolidation.

Acquisition des données

Données externes : Afin d'intégrer un maximum d'informations ayant un pouvoir de prédiction important au modèle de machine learning (d'apprentissage) qui sera mis en œuvre par la suite, les données externes suivantes ont été intégrées :

- Les données météorologiques : celles-ci intègrent chaque jour les données
 - D'ensoleillement,
 - De vent,
 - De couvertures nuageuses,
 - De taux de précipitations
 - Ainsi que la température maximale de la commune où se situe la préfecture ou la sous-préfecture la plus proche de la plage.
Nb : Dans le cas où la préfecture et la sous-préfecture se trouvent trop éloignées de la plage, ont été prises en compte les données météorologiques de la commune où se situe la plage. Ces données ont pu être récupérées toutes les heures.
- Les données calendaires : celles-ci intègrent
 - Les jours fériés,
 - Les vacances scolaires (avec zones hors été si besoin)
 - Jours de la semaine afin de détecter les comportements selon le jour de la semaine (week-end, mercredi...)
- Ont également été intégrées la date et l'heure afin de détecter les comportements selon les heures de la journée

Point à retenir :

- Ajout de données complémentaires pour intégrer le comportement des usagers, prise en compte des connaissances du GIP sur le comportement des usagers.

Segmentation

Caractérisation des plages

Dans un premier temps, l'idée était de créer une matrice permettant de caractériser l'ensemble des plages de Nouvelle-Aquitaine. Pour cela, un ensemble de variables a été repris de sources de données déjà existantes, les données open source ont été complétées, et d'autres variables ont été complétées par le GIP Littoral.

Une cinquantaine de variables ont été proposées afin de mettre en place une caractérisation des plages (Cf. Annexe 1).

Point à retenir :

- Identification des informations minimales permettant de regrouper les plages de manière qualitative.
- Utilisation des données existantes fournies par le GIP Littoral, issues du projet Données Plages¹¹ (disponibles sur <https://www.pigma.org/>) :

Nom variable	Définition
id_plage	identifiant de la plage
nom_plage	nom de la plage
commune	nom de la commune sur laquelle se situe la plage
insee_commune	numéro insee de la commune sur laquelle se situe la plage
plage_ambiance	ambiance de la plage : plage naturelle, plage urbaine ou aucun des deux
nature_baignade	nature de la baignade : océanique, lacustre, bassin de baignade, bassin d'Arcachon
grand_territoire	zone géographique sur laquelle se trouve la plage (Cote Basque, Nord Landes, Bassin d'Arcachon..)
stationnement	Véhicules pour lesquels le stationnement est possible (ex : voitures, vélos, camping-cars..)
equipements_services	équipements disponibles sur la plage (restaurant)
long	longitude
lat	latitude
insee_dep	numéro insee du département où se situe la plage
iframe_parking	plage équipée ou pas d'un capteur sur le parking
pression_touristique	plage soumise à pression touristique ou pas
payant	parking payant ou pas
pression_métropole	plage soumise à la pression d'une métropole proche ou pas
risque_incendie	plage exposée ou pas à des risques incendie en période de forte sécheresse

¹¹ <https://www.giplittoral.fr/ressources/donnees-sur-les-plages>

- Recherche de quelques sources de données complémentaires

A cet ensemble de données ont été ajoutées d'autres variables pertinentes pour l'étude de la fréquentation des plages.

Données open source :

Variable	Définition	Source	Priorisation
Urbanisation	Notion liée à la densité de bâtiments en arrière de la plage	https://www.datatourisme.fr/ontology/core#Accommodation	Forte
Nombre de lit	Nombre de lits total de la commune	https://www.insee.fr/fr/statistiques/5416748 https://www.insee.fr/fr/statistiques/2021703#dictionnaire	Forte
Population commune	Population commune littorale	https://www.insee.fr/fr/statistiques/zones/2011101?debut=0&q=population+des+communes	Faible
Proximité métropole	Distance à la métropole la plus proche	https://www.google.com/maps?authuser=0	Moyenne

Données externes : Données identifiées et complétées par le GIP Littoral au cours de l'étude.

Variable	Définition	Priorisation
Accessibilité piéton	Plage accessible ou pas à pied	Moyenne
Accessibilité vélo	Niveau d'accessibilité vélo (passable, bien, très bien)	Moyenne
Accessibilité voiture	Plage accessible en voiture ou pas	Moyenne
Capacité parking voiture	Nombre de places totales du ou des parkings voitures	Forte
Impact des marées	3 niveaux d'impact : impact sur la surface de plage disponible à marée haute / pas d'impact / pas de baignade possible à marée basse	Très forte
Plage équipée pour le comptage	Présence ou pas d'une caméra pour effectuer du comptage	Faible
Présence de commerce	Présence de commerces permanents	Forte

Clustering des plages (regroupement par caractéristiques)

Afin de ne pas équiper l'ensemble des plages de la région Nouvelle-Aquitaine, l'approche fut de caractériser les plages afin de regrouper celles qui sont similaires, et ainsi équiper qu'une portion représentative.

Un modèle de clustering a ainsi été appliqué afin de regrouper les plages similaires.

Point à retenir :

- Cette méthode de regroupement est basée sur des analyses mathématiques permettant d'éviter de prendre un biais basé sur des intuitions ou des raisonnements logiques. Cette technique permet de décrire la réalité de chaque plage de manière unique.
- K-means clustering

Le K-means clustering est un algorithme permettant de classer les données en plusieurs clusters. Pour fonctionner, il requiert d'indiquer au préalable le nombre de clusters dans lesquels classer les données. Il regroupe alors les données dans ces clusters en minimisant la distance au centroïde (point au centre du cluster) pour chaque donnée. En d'autres termes, si nous avons deux clusters, pour chaque donnée l'algorithme K-means calcule la distance entre cette donnée et le centre des deux clusters, avant d'attribuer la donnée dans un cluster en fonction de la distance la plus courte. L'algorithme fonctionne de manière similaire lorsque l'on augmente le nombre de clusters, en calculant la distance donnée/centre du cluster pour chaque centre de cluster.

- Méthodes Elbow et silhouette score

Pour connaître le choix du nombre de clusters, les méthodes « Elbow » et « Silhouette Score » ont été appliquées. La **méthode « Elbow »** consiste à calculer pour chaque nombre de clusters, le pourcentage de variance expliqué. Plus le nombre de clusters est grand, plus le pourcentage de variance expliqué sera grand (i.e moins d'information sera perdue). Le critère d'Elbow (=du coude) consiste à choisir le nombre de clusters à partir duquel, l'ajout de cluster n'ajoute plus que très peu de variance expliquée. A partir de ce critère, l'ajout de cluster ne sera plus aussi "rentable" qu'avant.

La **méthode du « silhouette score »** consiste à calculer un coefficient, pour chaque nombre de clusters, mesurant la similarité entre les points d'un même cluster et la dissimilarité entre les points de deux clusters différents. Plus ce coefficient est élevé, plus le clustering aura construit des groupes de points proches entre eux et distincts de ceux des autres clusters (*voir le détail du calcul du coefficient de silhouette score en annexe 4*).

Consolidation

Données d'entrée du modèle d'estimation : Le modèle d'estimation de fréquentation des plages est ainsi basé sur les variables suivantes :

- Les données de caractérisation des plages
- Le résultat du clustering calculé selon la méthode décrite précédemment
- Les données météorologiques du jour :
 - Ensoleillement
 - Taux de précipitations
 - Température maximale
- Le calendrier :
 - Jours fériés
 - Vacances scolaires
- La date :
 - Jour de la semaine
 - Mois de l'année

Afin de répondre le plus précisément possible sur l'estimation de la fréquentation des plages, plusieurs modèles de machine learning ont été testés pour identifier le plus performant.

Pour cela, les modèles suivants ont été testés :

- Gradient boosting
- Random forest¹²
- Combinaison des deux modèles précédents (gradient boosting et random forest)

L'importance de chacune des variables en entrée du modèle a pu être évaluée après l'application du modèle de machine learning.

Triangulation des données : Afin de valider certaines variables telles que la pression touristique et la pression métropole, de nouvelles variables ont été intégrées dans la matrice de caractérisation des plages, telles que la proximité de la plage à une métropole/agglomération, la population de la commune littorale dans laquelle se situe la plage, la présence de commerces ou encore le nombre de lits dans une distance de 10 minutes à pied de la plage.

Afin d'étendre également les données disponibles sur l'ancienne région Aquitaine à la Charente-Maritime, un exercice de triangulation a été mis en place.

Point à retenir :

- Des données supplémentaires non utilisées dans les méthodes initiales ont été intégrées afin de valider leur pouvoir prédictif et d'améliorer la qualité des résultats (= principe de la triangulation).

1.2.2 | Récapitulatif des robustesses

Ci-dessous, un récapitulatif comparatif des robustesses des méthodes appliquées en 2017 et en 2022.

Légende :

Très faible robustesse → 1 point
Faible robustesse → 2 points
Robustesse moyenne → 3 points
Forte robustesse → 4 points
Très forte robustesse → 5 points

Etapes	Robustesse 2017	Robustesse 2022
Acquisition des données : 1. Données externes 2. Campagne de fréquentation	Forte Forte	Forte Forte
Segmentation : 1. Caractérisation des plages	Faible	Forte

¹² Une Random Forest (ou Forêt d'arbres de décision en français) est une technique d'apprentissage automatique : c'est une méthode d'ensemble, c'est-à-dire qu'elle « met ensemble » ou combine des résultats pour obtenir un super résultat final. **Source :** <https://datascientest.com/random-forest-definition>

2. Clustering	Très faible	Forte
Consolidation : 1. Modèle statistique 2. Triangulation des données	Faible Très faible	Forte Moyenne
Score final	14	23

Ainsi, la mise en place de cette nouvelle méthode permet de passer d'une robustesse estimée à 14 en 2017 à une robustesse estimée à 23. L'amélioration de la confiance dans les prédictions est estimée à hauteur de 71% (passage d'un score de 14 à un score de 23).

Point à retenir :

- Validation du pouvoir prédictif de l'ajout de données externes, et regroupement par caractéristiques des plages, pour une amélioration significative des mesures de fréquentation d'une part, mais aussi de répliquabilité sur des plages sans mesures (ce qui n'était pas le cas dans les premières approches en 2017)

1.2.3 | Mise en œuvre du modèle Nouvelle-Aquitaine 2022

Acquisition des données

Données externes : Afin de mettre en place le premier modèle (clustering ou regroupement des plages de Nouvelle Aquitaine par similarités), la société Hupi a collecté dans un premier temps les données démographiques suivantes, à savoir le nombre de lits touristiques contenus sur chaque commune littorale ainsi que la population de la commune :

- Nombre de lits touristiques de la commune : cette donnée n'étant pas directement disponible, il a été nécessaire de récupérer indépendamment le nombre des différents lits touristiques, et effectuer par la suite une transformation linéaire formulée ci-dessous.

(Source : Insee : <https://www.insee.fr/fr/statistiques/5416748>
<https://www.insee.fr/fr/statistiques/2021703#dictionnaire>)

Formule utilisée pour calculer le nombre de lits par commune :

Nombre de lits = 2*(chambres d'hôtels) + 4*(emplacements camping) + (lits village vacances) + (lits auberge de jeunesse) + 5*(résidences secondaires)

- Population de la commune (Source : Insee, 2020)
<https://www.insee.fr/fr/statistiques/zones/2011101?debut=0&q=population+des+communes>)

Dans un second temps, afin d'étendre le modèle statistique de fréquentation à l'ensemble des plages de Nouvelle Aquitaine, les données météorologiques manquantes ont été récupérées au lien suivant via les coordonnées GPS¹³ :

<https://cds.climate.copernicus.eu/cdsapp#!/dataset/sis-agrometeorological-indicators?tab=overview>

¹³ Nb : La société HUPI a participé à des projets Copernicus et a ainsi pu réutiliser ces informations

Campagne de fréquentation : Les données de campagne de fréquentation utilisées sont les données de comptages effectuées en 2017 sur 18 plages du littoral aquitain.

Segmentation

Caractérisation des plages : La société Hupi a récupéré les données descriptives des plages¹⁴. Une grande partie des données a été récupérée à partir des données PIGMA existantes, et complétées en interne par le GIP Littoral pour mettre en œuvre et enrichir la clusterisation/segmentation des plages.

Traitement des données PIGMA :

- Indicatrice de présence d'un bassin de baignade
- Indicatrice de présence d'un parking pour camping-cars
- Indicatrice de présence d'un parking pour voitures
- Indicatrice de présence d'un parking pour vélos
- Nombre d'équipements présents sur la plage
- Indicatrice plage naturelle ou urbaine

Lorsque la mention "indicatrice" est précisée, cela signifie que la présence ou non de la caractéristique a été encodée en code binaire (0 et 1) : 1 si la caractéristique est présente pour la plage considérée, 0 dans le cas inverse.

Les données complétées en interne par le GIP Littoral sur l'ensemble des plages de Nouvelle Aquitaine sont les suivantes :

- Influence de la marée :
 - Niveau 1 : pas d'impact
 - Niveau 2 : impact sur la surface de plage disponible à marée haute
 - Niveau 3 : pas de baignade possible à marée basse
- Présence de métropole :
 - Non ⇒ 0
 - Oui ⇒ 0.5
 - Oui, métropole de Bordeaux ⇒ 1
- Accessibilité vélos (3 niveaux) :
 - Niveau 1 : accessibilité moyenne
 - Niveau 2 : bonne accessibilité
 - Niveau 3 : très bonne accessibilité
- Accessibilité piéton (3 niveaux) : distance en mètres depuis l'urbanisation et/ou le parking, jusqu'à la plage
 - Niveau 1 : > 600
 - Niveau 2 : 200-600
 - Niveau 3 : 0-200

En complément, les 2 données démographiques abordées précédemment ont été incluses (nombre de lits touristiques présents sur la commune et population de la commune).

Toutes ces données ont été **normalisées** afin de donner la même importance à chaque variable pour le clustering.

Normaliser une variable, c'est mettre la plus faible valeur prise par cette variable à 0 et la plus grande à 1. Les valeurs intermédiaires sont comprises entre 0 et 1, en conservant la 'taille' - la proportion - de l'écart entre les valeurs qui existait avant la normalisation (i.e, si avant normalisation

¹⁴ Projet Données Plages / Plateforme PIGMA

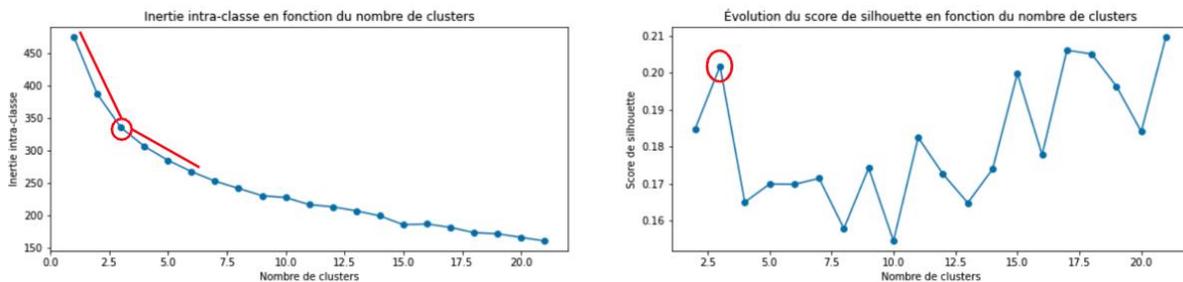
deux communes avaient un écart de population correspondant à 1000 habitants, ce qui correspond à $1/60$ du maximum d'habitants (60 000), alors après normalisation les communes ont un écart d'habitants de $1/60 = 0.017$). La normalisation sert à mettre toutes les données à la même échelle, sinon les algorithmes font des erreurs lors de la prise en compte des données. Exemple : la variable « Influence de la marée » est prise en compte selon des niveaux (1/2/3), alors que la variable de la population est de l'ordre du millier.

(Voir la formule de normalisation de données en annexe 4).

Clustering des plages : L'objectif du clustering fut de regrouper les plages similaires, et donc susceptibles d'avoir le même comportement (même "schéma" de fréquentations) pour donner des informations supplémentaires au modèle d'estimation. L'objectif fut alors d'identifier les variables qui différencient le plus les plages, afin de trouver la meilleure segmentation possible en fonction du nombre de groupes (clusters) choisis.

Nb : Après avoir effectué différents tests, il est apparu que les deux variables démographiques (nombre de lits de la commune et population de la commune) utilisées pour le clustering n'influaient pas lorsque le choix du nombre de clusters se limitait à 3 : les résultats du modèle donnaient exactement les 3 mêmes clusters avec l'intégration de ces variables ou non (à noter cependant que cette observation peut fluctuer avec le choix de prendre en compte un plus grand nombre de clusters). Ceci paraît cohérent car ces données ne sont pas des caractéristiques directes des plages. En revanche, ces variables sont un indicateur important dans l'étape de construction du modèle statistique de prévision de fréquentation des plages (à bien distinguer du modèle de clustering).

Pour déterminer le nombre de clusters nécessaires, les méthodes Elbow et Silhouette Score ont été utilisées (décrites précédemment).



Illustrations graphiques méthode Elbow à gauche et méthode Silhouette Score à droite.

La **méthode Elbow** repose sur une analyse visuelle. Il s'agit de repérer un "coude", i.e. une brusque décroissance sur le premier graphe. Comme indiqué en rouge sur la courbe, un tel coude semble se démarquer au 3ème point, ce qui enclenche à ne conserver que **3 clusters**.

Cependant, il n'y a pas ici de brusque décroissance flagrante, c'est pourquoi l'utilisation de la méthode du **Silhouette Score** est nécessaire pour confirmer le choix du nombre de clusters. Cette métrique est un critère d'ajustement des données. Plus celle-ci est élevée, plus les données sont bien séparées au sens "mathématique" du terme (i.e. en se basant sur la distance mathématique entre les différentes plages). Pour déterminer le bon nombre de clusters, il s'agit donc de prendre le nombre de clusters maximisant le silhouette score.

Ici, ce nombre optimal est encore **3**, si l'on considère les nombres de clusters inférieurs à 17.

Dans le cadre du modèle d'estimation de la fréquentation et de l'étendue des résultats de 2017 à 2022, le choix de séparer les plages en 3 clusters a été fait, car les données de comptage existantes ne sont disponibles que sur 18 plages (nécessaire d'avoir à minima une plage avec du comptage par cluster pour que le modèle fonctionne).

Un découpage plus fin en intégrant un nombre de clusters plus élevé pourrait être envisagé sur le long terme dans le cas de la mise en œuvre d'une campagne de fréquentation des plages, avec équipement de plages supplémentaires.

- Résultats du modèle de clustering :

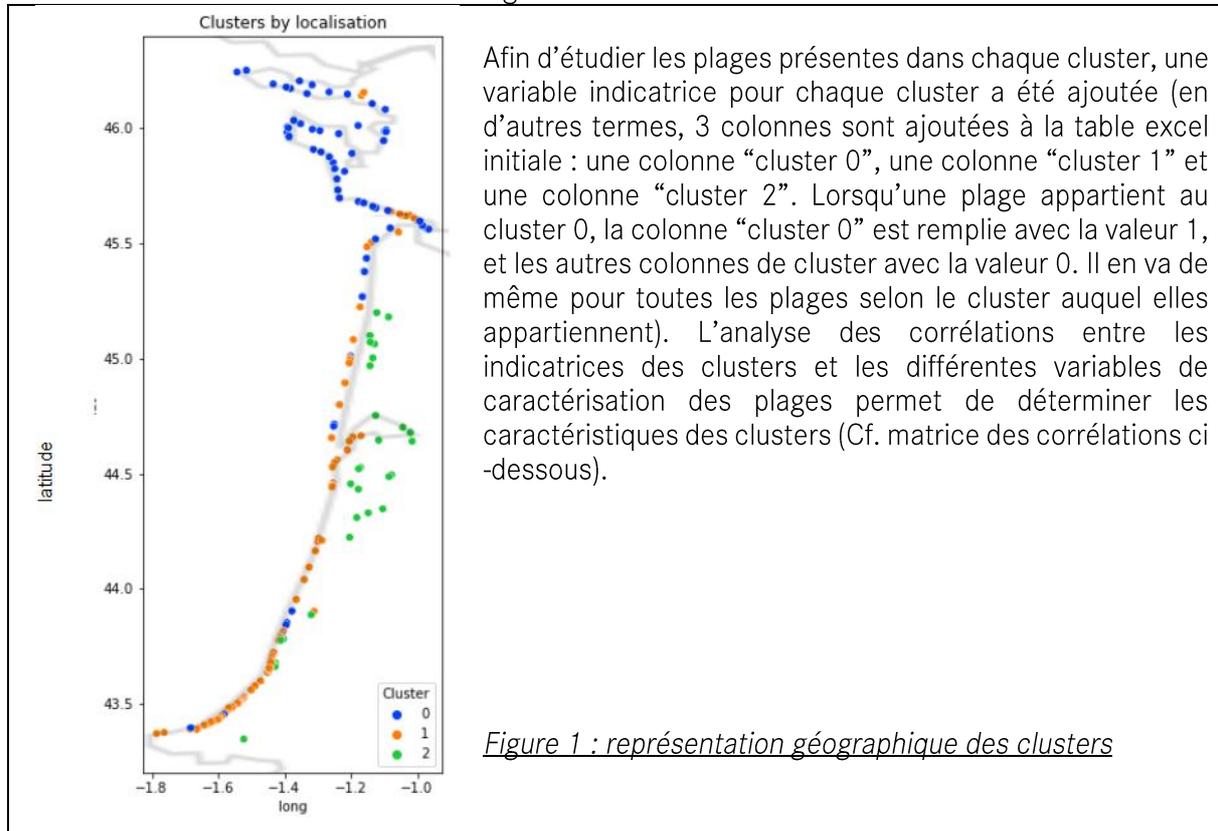


Figure 1 : représentation géographique des clusters

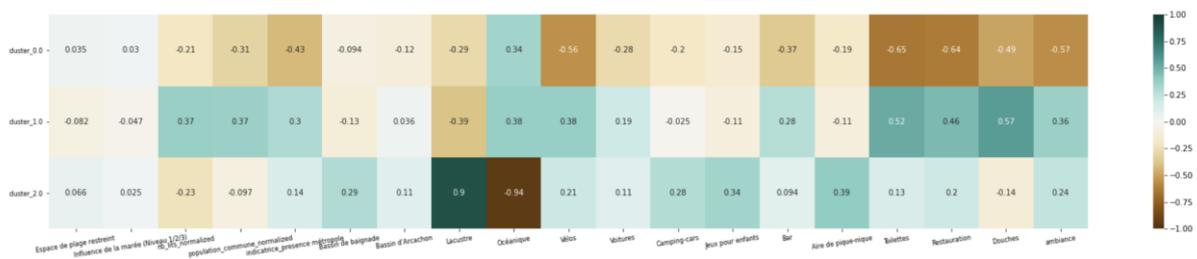
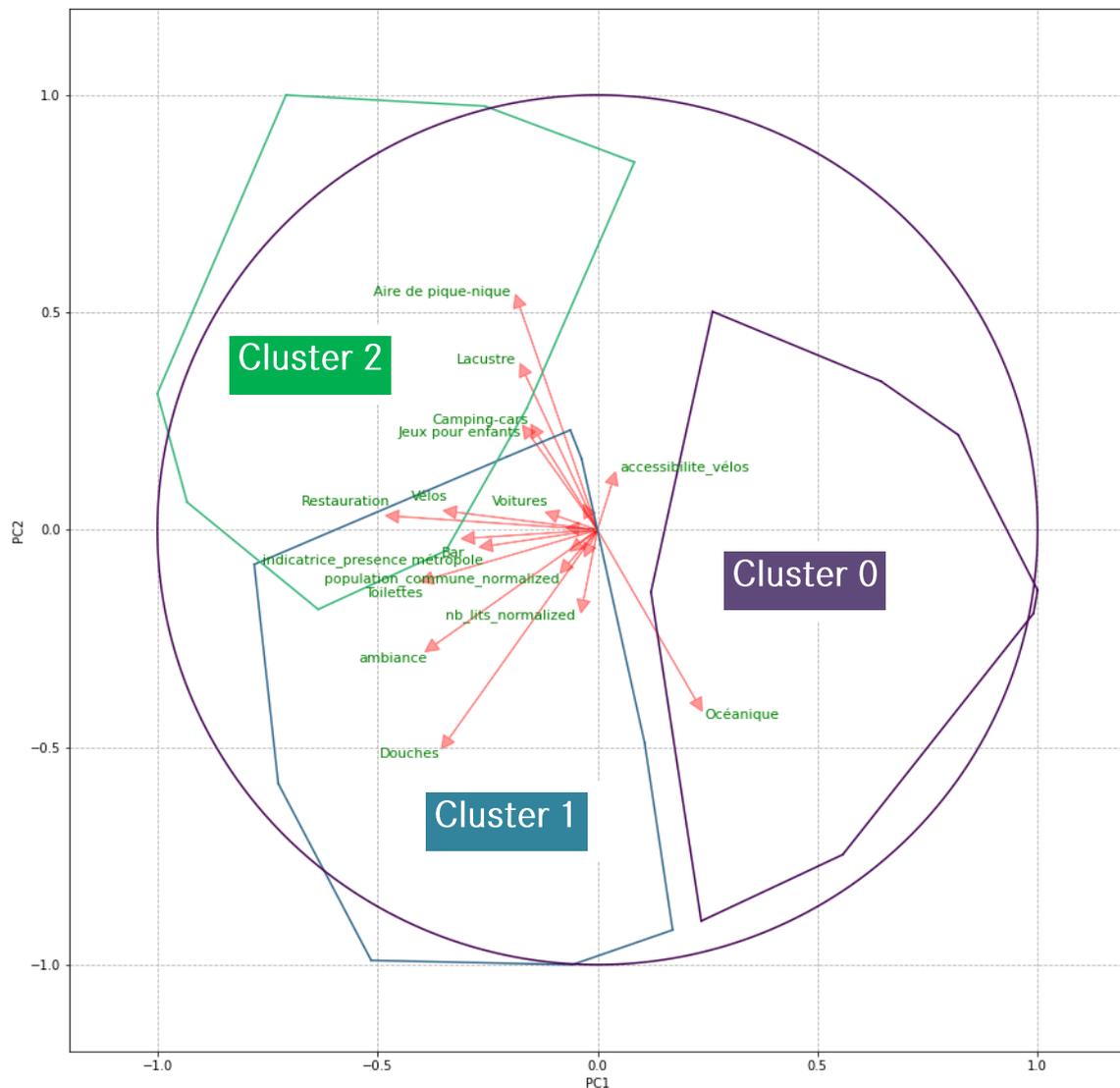


Figure 2 : Matrice des corrélations

Analyse : La corrélation varie entre 1 (vert foncé sur la matrice) et -1 (marron foncé). Une forte corrélation positive entre deux variables indique que si une variable prend une grande valeur pour une plage, alors l'autre prend généralement une grande valeur également. Une forte corrélation négative indique au contraire que si une variable prend une grande valeur pour une plage, alors l'autre prend généralement une petite valeur. Ici par exemple, la corrélation entre les indicatrices "cluster 2" et "plage océanique" est de -0.94. Cela signifie que lorsqu'une plage a une indicatrice d'une valeur de 0 pour "plage océanique" (qu'elle n'est donc pas océanique), alors il est très probable que son indicatrice "cluster 2" soit égale à 1. L'inverse est vrai aussi. Autrement dit, le cluster 2 contient essentiellement des plages non océaniques. Cela se vérifie sur la figure 1.

Cette analyse a été complétée d'une analyse en composante principale (ACP)¹⁵ et d'une projection des plages et des variables explicatives sur les premières dimensions :



Caractérisation des clusters : l'ensemble de ces analyses a permis de caractériser les clusters de la manière simplifiée suivante :

- **Cluster 0 :** essentiellement des plages océaniques, “naturelles”, relativement isolées et peu équipées.
- **Cluster 1 :** essentiellement des plages océaniques, bien équipées.
- **Cluster 2 :** essentiellement des plages lacustres et des bassins de baignade.

Description des clusters :

¹⁵ L'ACP permet de calculer des matrices pour projeter les variables dans un nouvel espace en utilisant une nouvelle matrice qui montre le degré de similarité entre les variables. <https://www.xlstat.com/fr/>

Le **cluster 0** regroupe des plages océaniques, généralement éloignées d'une grande métropole/agglomération, et avec une ambiance globalement naturelle, souvent peu équipées en restauration, bar, douches.



Plage de la Lette Blanche - Vielle Saint Girons (40)

Le **cluster 1** regroupe des plages océaniques et toutes celles du bassin d'Arcachon. Ce sont des plages plus souvent proches d'une métropole/agglomération que les autres plages, se trouvant dans une commune peuplée et/ou touristique. L'ambiance de la plage est donc généralement plus urbanisée que celles du cluster 0. La restauration et les douches sont des services qui y sont souvent proposés.



Plage centrale de Carcans (33)

Le **cluster 2** regroupe l'ensemble des plages lacustres, et uniquement ces plages. Celles-ci sont globalement plus équipées en jeux pour enfants que les autres plages et disposent plus souvent d'une aire de pique-nique. L'ambiance de ces plages est généralement décrite comme naturelle.



Plage du Moutchic - Lacanau (33)

Au-delà de ces caractéristiques principales, le clustering permet de répartir les plages de manière plus fine et sait par exemple dire si une plage océanique moyennement équipée est plus proche des plages du cluster 0 ou bien de celles du cluster 1.

Consolidation

Mise en œuvre du modèle statistique :

La prédiction de la fréquentation est un problème de régression qui repose sur l'élaboration d'un modèle d'apprentissage automatique supervisé. Ce modèle utilise un ensemble de **variables explicatives** pour déterminer la valeur d'une **variable réponse**. Ici, les **variables explicatives** sont les caractéristiques de la plage (ensemble des variables utilisées pour le clustering), les données météorologiques, le cluster auquel appartient la plage, ainsi que d'autres variables externes qui sont décrites plus bas. La **variable réponse**, celle que l'on cherche à prédire, est la donnée de fréquentation.

Les **données météorologiques** intégrées dans le modèle sont les suivantes : la température, le flux de radiation solaire et le flux de précipitations.

Nb : la couverture nuageuse également considérée au départ est une variable fortement corrélée au flux de radiation solaire, cette variable n'a donc pas été intégrée, et seul le flux de radiation solaire a été pris en compte.

Le modèle a été construit à partir d'un algorithme d'apprentissage : il a dû apprendre à utiliser les variables explicatives pour prédire au mieux la réponse. Il a utilisé pour cela des **données d'entraînement** (ici, les données de 2017 des 18 plages enregistrées pendant 4 mois) dont la réponse (la fréquentation) est connue.

Pour quantifier la performance du modèle et le valider, il lui a également fallu des **données de test** (non connues du modèle, car non fournies pour l'entraînement) dont la réponse est également connue : ces données ont permis de comparer les prédictions aux valeurs réelles attendues.

Il existe plusieurs algorithmes d'apprentissage qui permettent, à partir d'un jeu de données d'entraînement, de prédire les données réponses d'un nouveau jeu de données dont la réponse est inconnue. Les algorithmes d'apprentissage utilisés ont été les suivants : « **gradient boosting** » et « **forêt aléatoire** ».

- **Variables supplémentaires utilisées pour le modèle statistique de prédiction**

En complément des variables ayant permis la mise en œuvre du clustering, des variables importantes à prendre en compte afin de prédire la fréquentation des plages ont été ajoutées :

- Température maximum de la journée
- Flux de radiation solaire
- Flux de précipitations
- Indicatrice de vacances scolaires
- Numéro du mois
- Numéro de la semaine
- Moyenne glissante des températures maximums sur 5 jours
- Différence entre la température maximum du jour et celle des 5 jours précédents
- Indicatrices pour chacun des 7 jours de la semaine
- Indicatrice de jour férié

- **Séparation des données d'entraînement et des données de test**

L'objectif de la séparation des données en jeu d'apprentissage et jeu de test fut d'évaluer les performances du modèle de façon "neutre". L'évaluation du modèle avec des données lui ayant servi pour s'entraîner aurait favorisé le sur-apprentissage et donné un modèle très flexible (s'ajustant très bien aux données d'entraînement) mais aussi trop variable et non adapté à la prédiction.

Deux divisions ont été envisagées :

- Sélectionner aléatoirement 20% des données de 2017.
- Sélectionner une plage parmi les 18 correspondant aux données de test.

Dans le cadre de l'étendue des résultats de fréquentation aux plages sans données de comptage, il est apparu plus intéressant d'utiliser la seconde possibilité. Ainsi le modèle statistique appliqué s'est entraîné sur 17 plages, puis les performances du modèle ont été testées et évaluées sur les données d'une nouvelle plage qu'il ne connaissait pas.

La plage de test utilisée est la plage du Gressier sur la commune du Porge (Gironde). C'est une plage appartenant au cluster 1 (contenant 80 des 180 plages du littoral néo-aquitain), océanique et bien équipée. Les données de test ont donc été les 113 données (pour les 113 jours de relevés) de cette plage. Les données d'entraînement étaient celles des 17 autres plages.

- **Résultats obtenus sur la plage de test**

Le tableau ci-dessous récapitule les erreurs obtenues pour chaque modèle, selon les différentes métriques. Les résultats ont également été combinés, en faisant la moyenne obtenue pour les deux autres modèles.

	Forêt aléatoire	Gradient boosting	Combinaison des 2 modèles
R ²	0.63	0.73	0.69
RMSE	640	544	584
MAE	550	369	4040

Le modèle de **gradient boosting** s'est révélé être le meilleur au sens du R², du RMSE et du MAE. Ces derniers indicateurs signifient que le modèle gradient boosting est en moyenne plus proche de la réalité, et que ses prédictions sont plus "équilibrées" que celles des autres modèles. L'indice priorisé, quelle que soit la situation, est le RMSE car celui-ci pénalise davantage les modèles ayant des résultats plus aberrants. Compte tenu des résultats ci-contre, le modèle qui a été retenu pour la suite est donc le modèle de boosting.

La figure 3 ci-dessous illustre les performances des modèles sur les données de la plage test. En bleu, la fréquentation réelle de la plage test au cours des 113 jours est représentée. On distingue la période des vacances scolaires qui se caractérise par une fréquentation globalement plus élevée pendant une cinquantaine de jours. En orange, jaune et vert, ont été superposées les prédictions effectuées par les différents modèles.

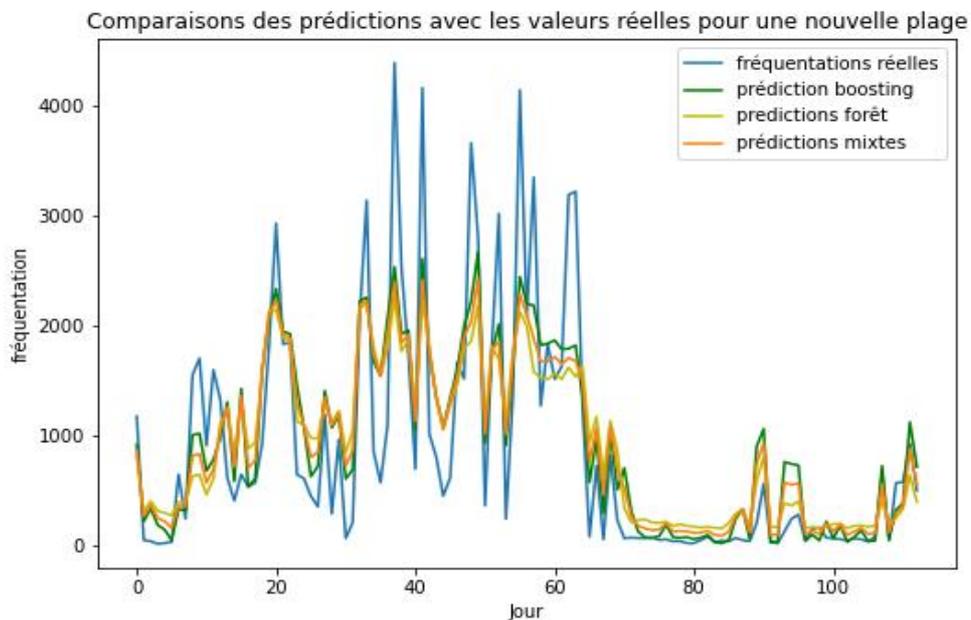


Figure 3 : Comparaison des prédictions avec les valeurs réelles pour la plage test

Le tableau ci-dessous affiche les écarts par mois de prédictions entre les fréquentations réelles (données de test) et celles prédites par le modèle combiné. Les calculs ont été faits en considérant 1 mois = 30 jours. Le dernier mois ne contient que 23 jours, en raison du nombre de données (113 jours de relevés) et n'est donc pas présenté dans le tableau.

Mois	26 juin - 25 juillet	26 juillet - 24 août	25 août - 23 septembre
Fréquentation réelle	25767	53355	13714
Fréquentation prédite	28649	52633	15468
Erreur de prédiction absolue	2882	722	1754
Erreur de prédiction relative	11,1%	1,4%	12,8%

Remarque : statistiquement, ces résultats ont été jugés satisfaisants par la société Hupi pour la mise en œuvre d'un modèle d'estimation fiable.

L'importance de chaque variable utilisée dans le modèle a pu être étudiée (Cf. Figure 4). Toutes les variables qui n'apparaissent pas ont une importance nulle. L'importance est une mesure de la contribution des variables à la réduction de l'erreur du modèle (pourcentage de réduction totale de la fonction de perte lors de leur séparation dans les arbres du modèle). L'ensemble des importances somme donc à 1. Cette mesure permet de déterminer quelles variables sont les plus influentes quant à la fréquentation des plages.

```

Variable: nb_lits_normalized      Importance: 0.2
Variable: normalized_solar_rad  Importance: 0.18
Variable: indic_vacances        Importance: 0.17
Variable: normalized_temperature Importance: 0.06
Variable: num_semaine           Importance: 0.06
Variable: cluster_1             Importance: 0.04
Variable: moyenne_glissante5j   Importance: 0.04
Variable: nombre_equipements    Importance: 0.04
Variable: population_commune_normalized Importance: 0.03
Variable: cluster_2             Importance: 0.03
Variable: diff_temp_moyenne5j   Importance: 0.03
Variable: Influence de la marée (Niveau 1/2/3) Importance: 0.02
Variable: Océanique             Importance: 0.02
Variable: normalized_precipitations Importance: 0.02
Variable: indicatrice_presence métropole Importance: 0.01
Variable: ambiance              Importance: 0.01
Variable: accessibilite_piétons  Importance: 0.01
Variable: accessibilite_vélos   Importance: 0.01
Variable: Dimanche              Importance: 0.01
Variable: Month                 Importance: 0.01

```

Figure 4 : Liste des variables utilisées dans le modèle et mesure de l'importance

Points à retenir :

- Les variables ajoutées suite à l'analyse de l'étude de 2017 sont celles qui influencent le plus le modèle.
- Les variables qui influencent davantage le modèle sont celles du **contexte de la journée (ensoleillement et vacances)** et non des caractéristiques de la plage, à l'exception du nombre de lits dans la commune.

1.2.4 | Résultats : Étendue des données de 2017 au périmètre de Nouvelle-Aquitaine



Vue aérienne du Nord Médoc et du Royannais

Données météorologiques utilisées

Afin d'étendre les résultats à l'ensemble des plages de Nouvelle-Aquitaine, il a fallu prendre des décisions quant aux données météorologiques à utiliser. Pour les plages de Charente-Maritime, le choix a été fait de récupérer les données météorologiques de la ville de Rochefort, centrale dans le département et proche de la côte. Les données ont été récupérées via les données Copernicus :

<https://cds.climate.copernicus.eu/cdsapp#!/dataset/sis-agrometeorological-indicators?tab=overview>

Pour les plages de l'ancienne région Aquitaine (Pyrénées-Atlantiques, Landes et Gironde), le même raisonnement que celui pris en 2017 a été suivi, sur lequel se base le nouveau modèle de prédiction de fréquentation des plages. Ainsi, les mêmes données de températures ont été utilisées pour les communes suivantes :

- Hendaye et Saint Jean de Luz
- Bidart, Biarritz et Ondres
- Seignosse
- Biscarrosse, Sanguinet, Gujan Mestras, Lege Cap Ferret, Le Porge, Lacanau, Carcans, Hourtin, Naujac sur Mer, Vendays Montalivet, Grayan et l'Hopital et Soulac sur Mer

Les mêmes données météorologiques ont été utilisées pour les communes suivantes :

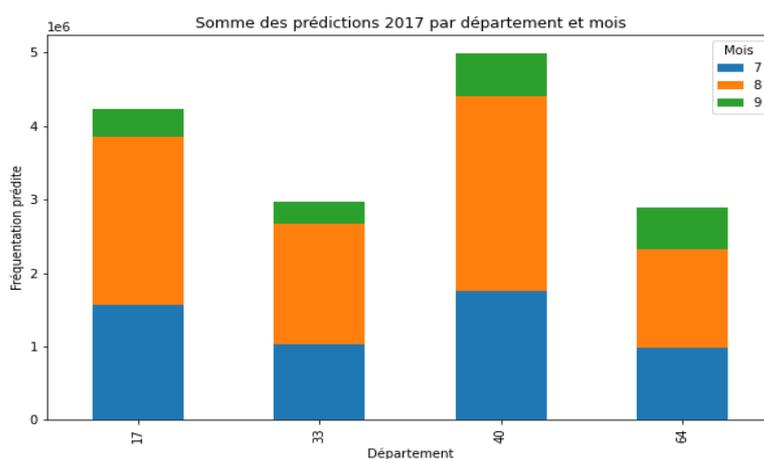
- L'ensemble des communes allant de Hendaye à Saint Jean de Luz
- L'ensemble des communes allant de Bidart à Labenne
- L'ensemble des communes allant de Capbreton à Vielle Saint Girons
- L'ensemble des communes allant de Lit et Mixe au Verdon sur Mer

Résultats obtenus

En premier lieu, le modèle de gradient boosting a été étendu aux 184 plages (incluant le périmètre de la Charente Maritime) avec pour objectif d'obtenir une estimation de la fréquentation en 2017. Les résultats sont représentés sous la forme du tableau et des graphiques suivants :

	Prédictions des fréquentations par mois et par département (en milliers) - 2017			
	Juillet	Août	Septembre	Total
Plages du littoral de Gironde (33)	1028	1651	300	2979
Plages du littoral des Landes (40)	1756	2651	588	4995
Plages du littoral des Pyrénées Atlantiques (64)	982	1346	558	2886
Plages du littoral de Charente Maritime (17)	1560	2286	381	4227
Total plages du littoral de Nouvelle Aquitaine	5326	7934	1827	15087

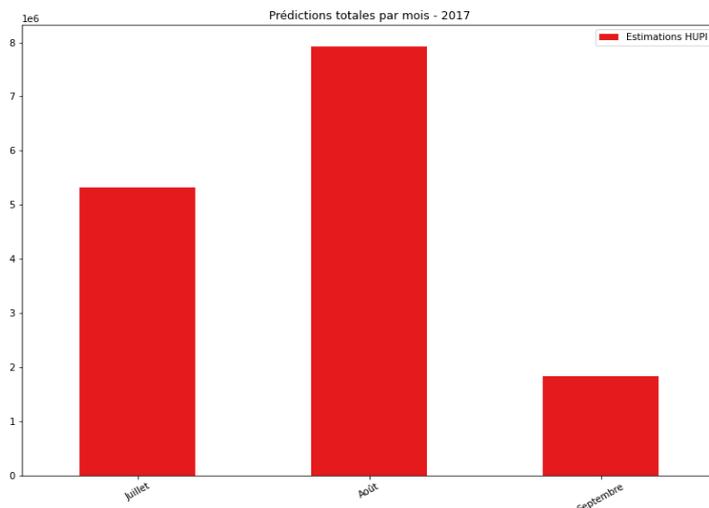
Pour rappel, l'estimation réalisée en 2017 avait permis d'obtenir le chiffre de 10 Millions de visites au total en saison estivale à l'échelle de l'ex-Aquitaine. La présente estimation comptabilise ce nombre de visites à 10.8 Millions de visites en ex-Aquitaine, et ajoute la part de visites du département de Charente Maritime, ramenant ce chiffre à 15 Millions à l'échelle des quatre départements en 2017.



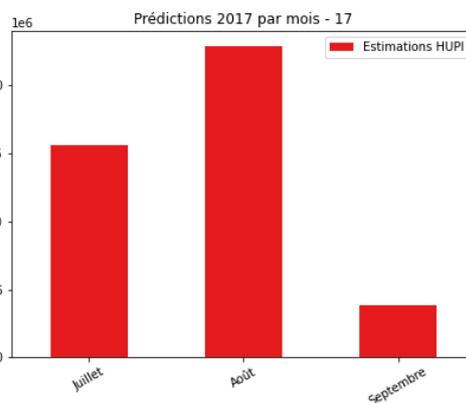
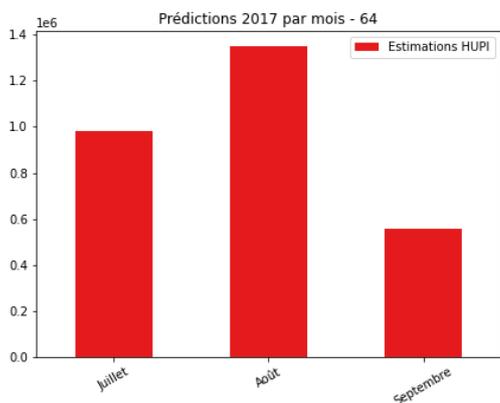
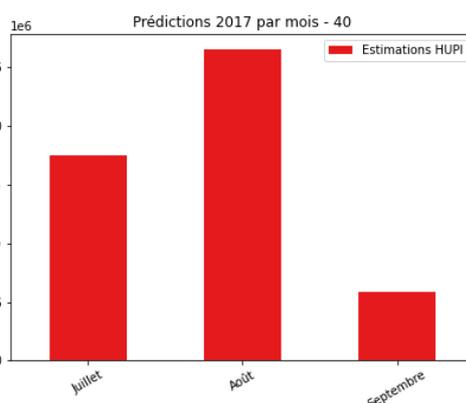
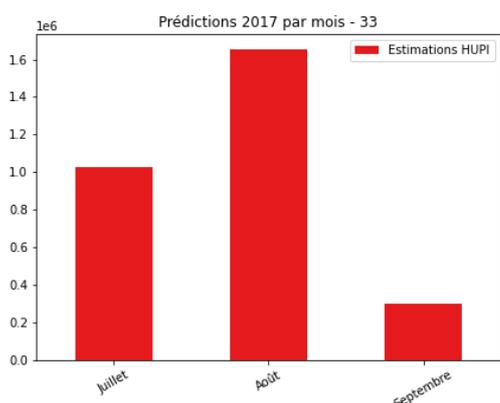
Afin d'interpréter correctement ces résultats, il est important de considérer en parallèle les caractéristiques côtières des départements. Celles-ci sont renseignées dans le tableau ci-dessous :

Département	Linéaire côtier (km)	Nombre de plages surveillées	Population	Capacité d'accueil touristique du département
33	201	42	1 636 391	529 146
40	107	55	418 122	395 724
64	35	31	687 240	339 168
17	306	56	655 709	653 540

Zoom sur les prédictions globales par mois en 2017, incluant le périmètre de la Charente Maritime :



Sans surprise, le mois d'août voit les plus grandes prédictions de fréquentation, suivi du mois de juillet puis du mois de septembre. En rentrant dans le **détail par mois et par département**, les graphiques suivants sont obtenus :



La tendance est la même dans chaque département. Cependant, l'effet du mois est plus ou moins marqué suivant les départements : par exemple la diminution de la fréquentation entre août et septembre est beaucoup moins marquée dans les Pyrénées Atlantiques (baisse de 59%) qu'en Charente Maritime (baisse de 83%).

Ci-dessous, le détail les évolutions des fréquentations prédites d'un mois à l'autre :

Territoire	Évolution de juillet à août	Évolution d'août à septembre
Gironde (33)	+61%	-82%
Landes (40)	+51%	-78%
Pyrénées-Atlantiques (64)	+37%	-59%
Charente Maritime (17)	+47%	-83%
Région Nouvelle Aquitaine	+49%	-77%

1.2.5 | Résultats : Étendue des données de fréquentation à 2022 en Nouvelle-Aquitaine

Données externes utilisées

L'extension des résultats de fréquentations des plages de 2017 à l'année 2022 s'est faite sur la base de **données touristiques de l'INSEE**¹⁶. La transformation s'est faite département par département, en se basant sur l'évolution du nombre de nuitées touristiques dans les hôtels et les campings entre 2017 et 2022.

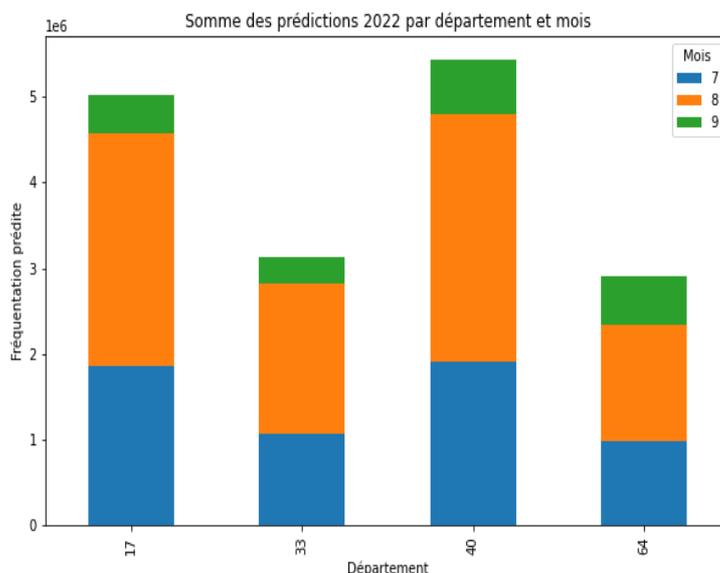
Les données récupérées ont reflété une progression du tourisme de 18,7% en Charente Maritime, de 0,5% en Pyrénées Atlantiques, de 8,8% dans les Landes et de 5% en Gironde. La même transformation a donc été appliquée à la prédiction de la fréquentation mensuelle des plages de 2017 en fonction du département dans lequel se trouvait la plage. Des **prédictions agrégées par mois et par département pour 2022** ont ainsi été obtenues.

Prédictions de la fréquentation des plages de Nouvelle-Aquitaine en 2022

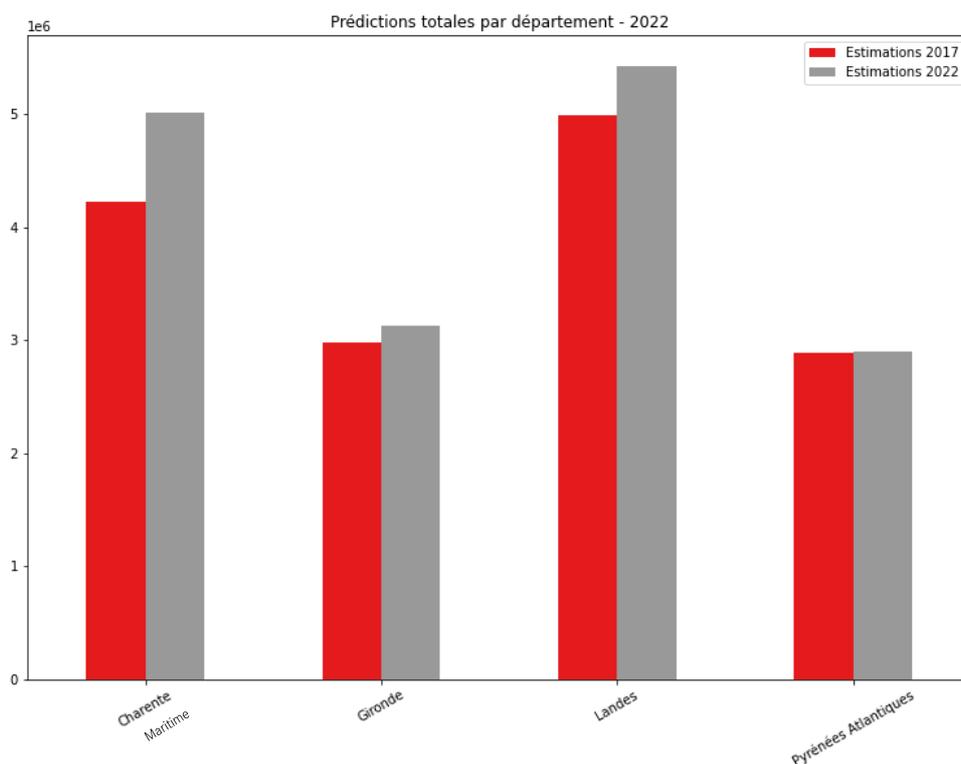
Une fois les transformations mathématiques réalisées, les chiffres suivants pour 2022 ont été obtenus :

	Prédictions des fréquentations par mois et par département (en milliers) - 2022			
	Juillet	Août	Septembre	Total
Plages du littoral de Gironde (33)	1079	1733	315	3127
Plages du littoral des Landes (40)	1910	2883	640	5433
Plages du littoral des Pyrénées Atlantiques (64)	987	1353	560	2900
Plages du littoral de Charente Maritime (17)	1852	2714	452	5018
Total plages du littoral de Nouvelle Aquitaine	5828	8683	1967	16478

¹⁶ <https://www.insee.fr/fr/statistiques/7342884?sommaire=7343353#consulter>
<https://www.insee.fr/fr/statistiques/3544028?sommaire=3550189&q=tourisme+nouvelle+aquitaine#tableau-figure1>



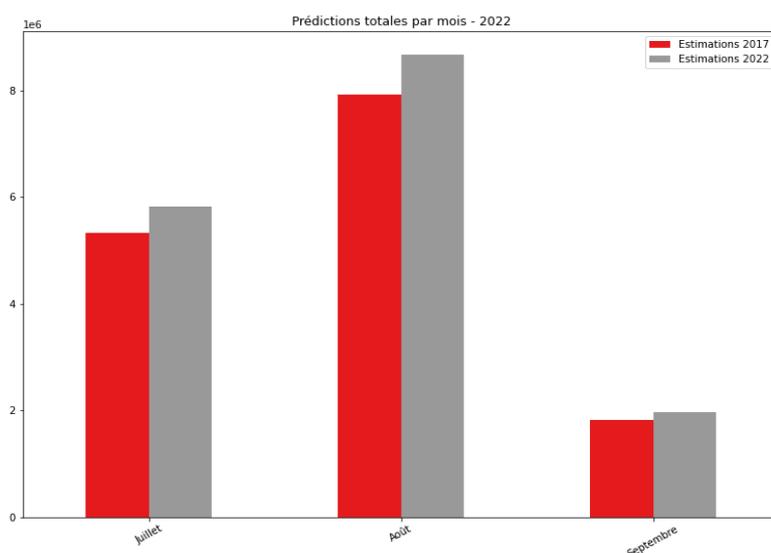
Dans l'ensemble, et en vertu de la tendance globale du tourisme français, les résultats sont donc en général plus élevés que pour 2017. Malgré les évolutions bien différentes suivant les départements, les tendances sont conservées : le 17 et le 40 restent les départements les plus fréquentés, suivis par le 64 et le 33.



Zoom sur l'évolution par département :

Conformément aux pourcentages d'évolution recueillis, l'augmentation est davantage marquée en Charente maritime et dans les Landes, qui étaient déjà les deux départements de la région les plus fréquentés en 2017. Par conséquent, l'augmentation en termes de fréquentation brute est encore plus marquée, ainsi que les différences entre départements.

Zoom sur l'évolution mensuelle globale :



L'évolution est très peu marquée en septembre. Cela s'explique par le fait que les Pyrénées Atlantiques est l'un des départements qui accueille le plus de touristes en septembre (derrière les Landes), or c'est aussi le département qui a l'évolution de fréquentation la plus faible entre 2017 et 2022, selon les données de l'INSEE.

Au contraire, en juillet et août, l'évolution est plus marquée : environ 1M de touristes de plus ont été estimés sur ces deux mois.

Points à retenir :

- Pas d'évolution majeure de la fréquentation des plages entre les saisons estivales 2017 et 2022 (+ 1.4 M environ, ce qui représente environ 8 % d'augmentation).
- Importance de la part de fréquentation des plages de Charente-Maritime au sein de la fréquentation régionale.

Couplée à une réalité touristique observée, les prédictions sans mesures terrains sur 2022 fournissent des résultats très qualitatifs. Ces informations peuvent être prises en compte dans une stratégie d'actualisation de l'estimation de la fréquentation régionale.

Partie 2 | Proposition de relance d'une campagne de fréquentation

Cette partie du rapport détaille la méthodologie préconisée par la société Hupi pour la mise en œuvre d'une campagne de fréquentation des plages à l'échelle de la Nouvelle-Aquitaine.

2.1 | Proposition méthodologique

S'il a été possible d'étendre les résultats de 2017 à 2022, pour disposer de données cohérentes dans le temps et afin de pouvoir détecter les évolutions, HUPI propose la méthodologie suivante de mise en place d'une campagne de fréquentation :

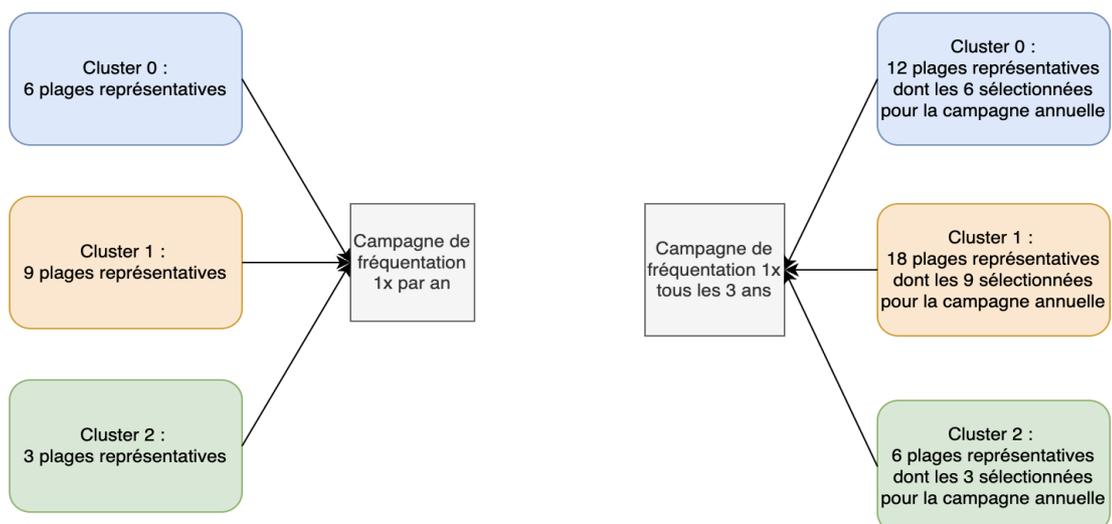
Deux types de campagnes de fréquentation complémentaires :

- Une première qui aurait lieu tous les 3 ans et qui consisterait à équiper entre 20% et 30% de plages par cluster pour permettre aux modèles de rééquilibrer les tendances.
- Une seconde qui serait réalisée tous les ans qui ne consisterait qu'à équiper les plages les plus représentatives des différentes tendances. Pour ces plages, il suffirait de sélectionner entre 10% et 15% des plages de chaque cluster. Ce type de campagne permettrait de détecter les tendances globales et ainsi de vérifier la cohérence avec les prédictions faites.

Les plages stratégiques à prendre en compte pour ces deux types de campagne seraient définies à l'aide des résultats du regroupement des plages réalisé dans le cadre de la phase 1 de l'étude (clustering).

Les plages types à équiper tous les ans seraient donc établies en début de campagne, à l'aide du clustering, puis complétées tous les 3 ans par d'autres plages types. Il pourrait être intéressant de ne pas équiper exactement les mêmes plages tous les 3 ans afin d'avoir le maximum de diversité possible dans les données d'entraînement du modèle.

Voici un schéma récapitulatif de la méthodologie proposée :



2.1.1 | Stratégie de déploiement

Pour rappel, la commande initiale du GIP et de ses partenaires fixait pour objectif de disposer :

- D'un système pérenne pour réaliser un suivi régulier,
- D'une méthode de calcul fixe dans le temps,
- D'une politique de suivi de la fréquentation du littoral exploitable.

Pour y répondre, la solution proposée chercherait à :

- Couvrir l'ensemble des contraintes des différents sites,
- Être une solution autonome sans intervention humaine,
- Intégrer des informations environnementales et contextuelles pour analyser les résultats (météo, température, pression touristique, ...),
- Proposer des coûts de maintenance réduits.

La stratégie reposerait ainsi sur une solution complète intégrant la solution matérielle et logicielle de bout en bout.

Zoom sur la solution proposée

Une solution tout intégrée, incluant le matériel et le software.



Un système de comptage à vision artificielle :

1. Le système de vision capture, analyse et détruit les images en temps réel : stocks (ex. parking) ou flux (ex. route d'accès).
2. L'algorithme d'Intelligence Artificielle embarqué extrait en temps réel des images les données de trafic (personnes, volume de voitures, bus, scooters...).
3. Le système de communication transmet en continu uniquement les données de trafic (volume) via réseau WIFI, LoraWan ou 3G/4G.
4. Le système utilise des systèmes d'alimentation électrique existants ou des systèmes autonomes fournis (alimentés par panneaux solaires).
5. Installation et formation : un soutien opérationnel pour l'installation et la maintenance à long terme est fourni par le prestataire.

Un système qui se fixe facilement sur différents mats de supports :



2.1.2 | Chronologie de déploiement



- **Identification des sites**

L'objectif de l'identification des sites serait de s'assurer de la maximisation de la couverture minimale nécessaire pour obtenir une évaluation de la fréquentation de l'ensemble du Littoral Néo-Aquitain.

Les sites seraient soumis à une grille d'évaluation prenant en compte les éléments suivants :

- Une sélection équilibrée de sites respectant les caractérisations des études et les distributions.
- Une évaluation des besoins des collectivités, ainsi que leurs motivations à intégrer le projet.
- L'aménagement global des sites (électricité, poste de secours, ...)

- **Préparation des sites**

L'objectif de la préparation des sites serait de s'assurer que tout le matériel et la procédure soit réalisée en amont de l'installation sur le site.

Les livrables et documents fournis comporteraient à minima :

- Le mode opératoire à suivre en fonction du site (incluant la zone à surveiller pour le comptage)
- La liste du matériel fourni
- La liste du matériel à fournir par le site (électricité et point d'accroche du système)
- Outils de gestion du projet

- **Déploiement**

La phase de déploiement correspondrait à la mise en fonctionnement du système incluant les opérations suivantes :

- Installation du système sur le site
 - Evaluation de la qualité de pose (robustesse, propreté)
 - Calibrage et orientation du système pour couvrir une zone prédéfinie
 - Tests de remontée des informations (simulation de coupure électrique, évaluation de la qualité du réseau, évaluation des comptages réalisés)
- **Maintenance**

Les opérations de maintenance incluraient des opérations liées principalement à des problèmes pouvant apparaître sur des systèmes (dégradation volontaire ou involontaire, dysfonctionnement, etc.).

2.2 | Estimation technique et financière

Dans le cadre de cette proposition méthodologique, la société Hupi a élaboré un **outil support pour estimer le chiffrage de l'opération, ainsi que les moyens humains à engager.**

2.2.1 | Support de Chiffrage

Un outil paramétrable prenant en compte les contraintes du terrain ainsi que les éléments constitutifs de l'étude, et permettant d'adapter un dispositif complet pour réaliser une étude de fréquentation sur l'ensemble du Littoral est proposée. Dans le cadre de cet outil, deux solutions financières sont proposées : une solution **d'achat** et une solution de **location** matérielle et logicielle.

Comparaison des deux options :

La solution « achat » / investissement comprend de l'acquisition de matériel, impliquant par ailleurs la prise en compte l'obsolescence des équipements de comptage, ainsi qu'un temps de travail plus important du côté de la collectivité.

La solution de location, quant à elle, est garantie dans le temps et permet une plus grande souplesse dans la gestion des équipements installés, mais représente un coût plus onéreux.

- **Estimation prévisionnelle de la solution matérielle et logicielle « Achat »**

SERVICES	Détail	Fréquence	Tarif HT
Acquisition Systèmes Cyber- physiques de Comptage	Sur base de 79 systèmes répartis, incluant le logiciel dédié au comptage embarqué	Installation	134 805 €/HT
Abonnement	Gestion des informations Rendus graphiques (uniquement sur les informations remontées par les systèmes)	Mensuelle	1 000 €/HT/mois

Maintenance (correspondant uniquement au support en cas d'intervention et d'analyse)	Sur base de 79 systèmes répartis	Annuelle	1 975 €/HT/an
Garantie des équipements	La garantie est de 1 an à partir de la date de pose		Hors rachat garantie d'un « device »

- Estimation prévisionnelle d'une solution matérielle et logicielle « Locations »

Location sur 3 ans des équipements.

Le coût d'achat du matériel estimé est de 135 000 €/HT, une location annuelle est estimée à 59 000 €.

Les coûts de fonctionnement pour la solution « achat » représentent environ 14 000 €/HT par an (abonnement et maintenance).

Dans le cas de la solution « location », la maintenance n'est pas à prendre en compte car incluse dans la location, et donc les coûts de fonctionnement sont de 12 000 €/HT par an.

Une garantie est intégrée tout au long de la location permettant ainsi d'avoir un budget fixe sans risque dû à des pannes ou des remplacements de boîtiers.

2.2.2 | Pilotage et suivi du projet

Le déploiement de la solution proposée nécessiterait de prévoir du temps de suivi et de pilotage, via la mise en œuvre d'une équipe projet.

Une estimation des moyens humains à engager selon les différents acteurs a été réalisée :

Collectivités	GIP Littoral	Prestataire
36 j/h	44,5 j/h	50,5 j/h

Derrière cette répartition des moyens humains à engager, les missions prévues sont les suivantes :

- **Pour le GIP Littoral :** Suivi et coordination globale du projet (animation, communication, coordination collectivités/prestataire), accompagnement du prestataire dans le choix des sites à équiper, appui technique aux collectivités dans le suivi du projet et des temps de validation.
- **Pour le prestataire :** Appui au GIP sur la fourniture des supports de communication du projet, choix des sites à équiper, caractérisation/préparation des sites pour le déploiement des dispositifs (paramétrage des systèmes de comptages, fiches techniques installation), calibrage des dispositifs après installation, tests fonctionnement et remontée d'informations, fourniture de supports techniques pour la maintenance et l'entretien des dispositifs.
- **Pour les collectivités :** Installation des dispositifs et mise en fonctionnement, arbitrages et validations.

2.2.3 | Outil de simulation d'une campagne de fréquentation des plages

Afin de faciliter le travail de mise en œuvre d'une campagne de fréquentation des plages, la société Hupi a également mis à disposition du GIP un outil modulable (simulateur) permettant de paramétrer le cout et le dimensionnement du dispositif selon les besoins. Il contient notamment la liste complète des plages surveillées de Nouvelle-Aquitaine et des informations de contexte (ville/département), ainsi que le cluster auquel elles appartiennent selon la méthode qui a été définie dans l'étude. Ensuite, plusieurs critères modulables sont proposés permettant d'ajuster :

- La sélection des plages à équiper.
- Par plage sélectionnée, le type de campagne prévu (continue ou d'ajustement).
- Le nombre d'équipements à prévoir sur la plage.
- La précision d'un certain nombre de contraintes présentes sur sites (type d'alimentation électrique et de réseau fournis sur site, si un mat d'installation est à fournir, si l'installation requiert au préalable une étude sur le terrain).

Le remplissage de ces différentes rubriques permet ensuite d'obtenir le prix du dispositif, de l'installation, et des interventions de maintenance par plage équipée, ainsi que le total des plages.

Aperçu de l'outil de simulation :

										66	79	134 805,00 €	34 050,00 €	1 975,00 €			
Cher	Plages	Ville	Région	ressélec	Type de Campagne	#équipements	Energie	Réseau	Mat non disponible	Site non coté	Tarifs HT Plage équ	Installat	Intervention Mainte	GIP	Services E	HUPI	
1	Plage Centrale	SOUJAC SUR MER	33	X	Campagne ajustement	1	Alimentation électrique fournie ou 12 ou 24V (min 3A)	Réseau 4G	Mat à fournir unic Etude de fais		1495	550	74,75				
0	Plage Centrale	VENDAYS MONTALIV	33	X	Campagne continue	2	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		2877	550	143,85				
0	Plage Sud	VENDAYS MONTALIV	33	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
0	Plage du Pin Sec	NAUJAC SUR MER	33	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
2	Plage de Piquoyrot	HOURTIN	33	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage Centrale	HOURTIN	33	X	Campagne continue	2	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		2877	550	143,85				
2	Plage du Montaut	CARCANS	33	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
0	Plage Nord	LACANAU	33	X	Campagne ajustement	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage Centrale	LACANAU	33	X	Campagne continue	2	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		2877	550	143,85				
1	Plage Sud	LACANAU	33	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage Supér Sud	LACANAU	33	X	Campagne ajustement	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage du Gressier	LE PORGE	33	X	Campagne continue	2	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		2877	550	143,85				
1	Plage du Grand Crohot	LEGE CAP FERRET	33	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
0	Plage du Truc Vert	LEGE CAP FERRET	33	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
2	Basain de Baignade de Lanton	LANTON	33	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
2	Plage de La Humie	GUJAN MESTRAS	33	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage Fleurié	ARCACHON	33	X	Campagne continue	2	Alimentation électrique fournie ou 12 ou 24V (min 3A)	Réseau 4G	Mat à fournir et Etude sur Terr		3145	550	157,25				
1	Plage de la Corniche	LA TESTE DE BUCH	33	X	Campagne continue	1	Alimentation électrique fournie ou 12 ou 24V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1895	550	94,75				
1	Plage de la Salle Nord	LA TESTE DE BUCH	33	X	Campagne continue	2	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude de fais		2477	550	123,85				
1	Plage Centrale	BISCARROSSE	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude de fais		1527	550	76,35				
2	Plage de Navarrosse	BISCARROSSE	40	X	Campagne continue	2	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		2877	550	143,85				
2	Plage du Port de Maguide	BISCARROSSE	40	X	Campagne continue	2	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		2877	550	143,85				
1	Plage Nord	MMIZAN	40	X	Campagne continue	2	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude de fais		2477	550	123,85				
1	Saint Grons Plage	VIELLE SAINT GIRON	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage du Col Vert	VIELLE SAINT GIRON	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
0	Plage Centrale	MOUETS ET MAA	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
0	Plage des Chênes lièges	MOUETS ET MAA	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage centrale	MESSANGES	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
0	Plage Sud	MESSANGES	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
2	Plage du Lac Marin	VIEUX BOUCAU	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage du Péron	SEIGNOSSE	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage des Bourdaines	SEIGNOSSE	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
0	Plage de la Gravière	SOORTS HOSSEGOI	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
2	Plage du Rey	SOORTS HOSSEGOI	40	X	Campagne continue	2	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		2877	550	143,85				
0	Plage Blanche	SOORTS HOSSEGOI	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage Centrale	SOORTS HOSSEGOI	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage de La Pistie	CAPRETON	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage des Océandes	CAPRETON	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage Santocha	CAPRETON	40	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage de La Barte	ANJLET	64	X	Campagne ajustement	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude de fais		1527	550	76,35				
1	Plage des Cavaliers	ANJLET	64	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude de fais		1527	550	76,35				
1	Plage de la Petite Chambre d'Anour	ANJLET	64	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage des Sables d'Or	ANJLET	64	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
0	Plage du Pavillon royal	BDART	64	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				
1	Plage de Partermenta	BDART	64	X	Campagne continue	1	Alimentation électrique fournie 220V (min 3A)	Réseau 4G	Mat à fournir unic Etude sur Terr		1927	550	96,35				

Conclusion

La plus-value de la méthodologie proposée dans le cadre de ce rapport d'étude porte donc sur l'élaboration d'un modèle permettant d'envisager la mise en place d'une campagne de fréquentation des plages à la fois fiable, et sans suréquiper les sites de plage. Le travail réalisé constitue une optimisation du traitement des données existantes en matière d'actualisation des chiffres de la fréquentation régionale des plages, sans nouvelle production de données.

Cette étude a ainsi permis au GIP Littoral et ses partenaires de disposer de deux éléments clés répondant aux besoins formulés, identifiés dans le cadre de la demande initiale :

- **Une réévaluation en 2022 de la fréquentation des plages intégrant le périmètre régional de la Nouvelle-Aquitaine, sans mesures terrain**

Le travail effectué dans le cadre de la première partie de l'étude a permis à la fois d'actualiser l'estimation de la fréquentation régionale des plages en 2017, et d'étendre cette nouvelle estimation à 2022.

Ainsi, la fréquentation 2022 des plages de la région Nouvelle-Aquitaine en saison estivale est estimée à **16 478 000 visites**.

A l'échelle de **l'ex-Aquitaine**, la fréquentation est ainsi passée de **10 Millions** de visites estimées en 2017 dans le cadre de la première campagne de fréquentation, à **11,5 Millions** de visites estimées en 2022 à partir de la méthode actualisée.

La part de la fréquentation estimée en **Charente-Maritime** a ainsi eu un impact fort sur le chiffre régional, car elle constitue **30%** des visites sur le littoral régional en 2022.

L'augmentation de la fréquentation entre 2017 et 2022 est plus marquée en Charente maritime et dans les Landes, ce sont par ailleurs les deux départements de la région dont les plages sont les plus fréquentées.

Sur la totalité des plages des 4 départements, ce sont environ **1.2 Millions de visites** de plus estimées en **juillet et août**, et **140 000** de plus sur le mois de **septembre**, entre 2017 et 2022.

- **Une méthodologie complète et opérationnelle** permettant d'actionner à l'avenir, et selon le besoin, **la mise en œuvre d'une nouvelle campagne de fréquentation des plages**.

A l'issue de cette deuxième phase, le GIP Littoral et ses membres bénéficient à la fois d'une donnée actualisée sur la fréquentation des plages à l'échelle de la Nouvelle-Aquitaine, ainsi que d'une proposition technique et financière permettant d'ajuster et calibrer la mise en œuvre d'une campagne de fréquentation des plages selon les besoins. Ainsi, on estime aujourd'hui le budget de relance d'une campagne de fréquentation remplissant les objectifs initiaux à un investissement de **135 000€** et **des coûts de fonctionnement d'environ 14 000€ par an**.

La mise en œuvre d'une campagne de fréquentation représente ainsi des coûts partagés importants, nécessitant également la mobilisation d'une certaine ingénierie pour le suivi du projet.

La mission d'étude s'inscrit pleinement dans un contexte plus large de tourisme durable, dans lequel les territoires sont plus que jamais concernés par les problématiques de gestion des flux et de la fréquentation touristique. Le tourisme est en effet une grande force économique pour le littoral de Nouvelle-Aquitaine, qui voit en contrepartie certains de ses sites touristiques (en majorité les espaces littoraux) subir de forts pics de fréquentation, pouvant générer des impacts sur l'environnement, le cadre de vie des habitants, et conduire à une qualité d'accueil dégradée pour les visiteurs.

L'actualisation des chiffres de la fréquentation régionale des plages, ainsi que les réflexions autour de la mise en place d'une campagne régionale de mesure de la fréquentation des plages, se

retrouvent ainsi en cohérence avec le contexte général sur ces questions, qui aujourd'hui se concrétise notamment via l'Appel à Manifestation d'Intérêt lancé par l'Etat, dans le cadre de sa stratégie de gestion des flux touristiques (annoncée en juin 2023), et pour lequel certains territoires de Nouvelle-Aquitaine ont déjà manifesté leur intérêt.

Les collectivités s'inscrivent en outre dans une démarche qualitative concernant la connaissance des flux de fréquentation. Si le quantitatif permet de calibrer certains besoins liés au fonctionnement des sites, tels que les besoins en surveillance des plages par exemple, une meilleure connaissance qualitative des flux permettrait de répondre à des besoins en connaissance majeurs pour calibrer le fonctionnement et l'accueil sur les sites : par exemple, la mise en place d'alternatives à la voiture. Dans le cadre de la poursuite des réflexions s'inscrivant dans la continuité de cette étude, un volet qualitatif a vocation à être pleinement intégré.

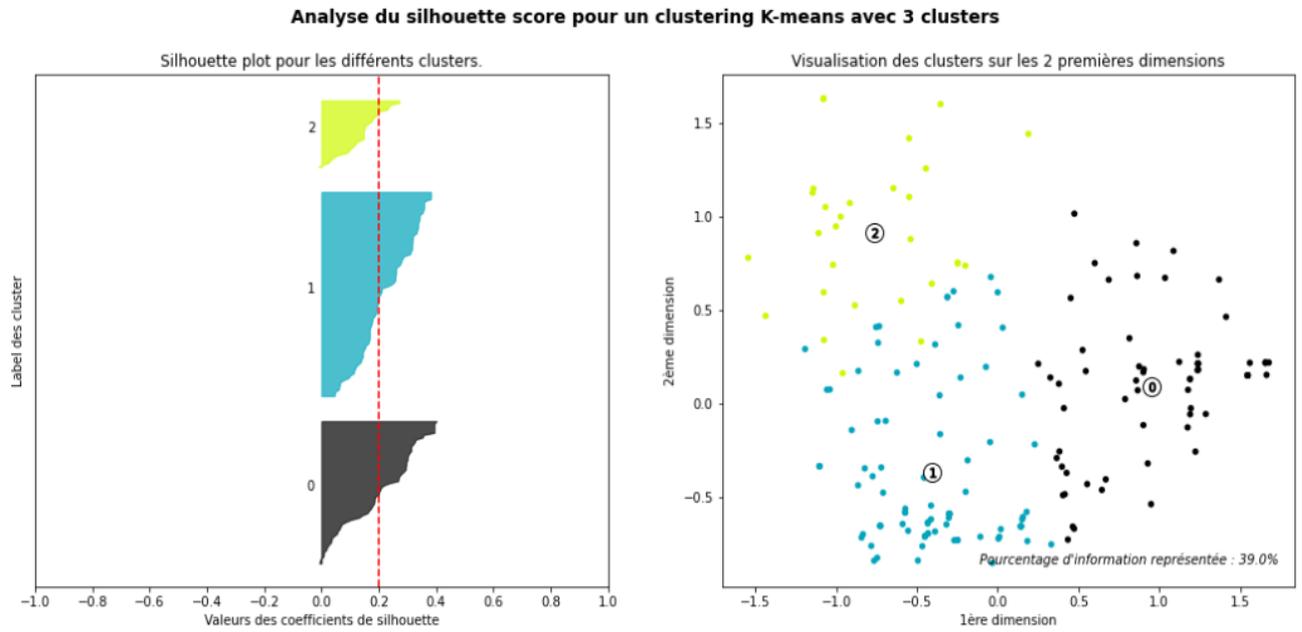
Annexes

Annexe 1 : Variables de caractérisation des plages proposées

Variable	Définition variable
nom_plage	nom de la plage
commune	nom de la commune sur laquelle se situe la plage
insee_commune	numéro insee de la commune sur laquelle se situe la plage
plage_ambiance	ambiance de la plage : plage naturelle, plage urbaine, ou aucun des deux
plage_caracteristiques	caractéristiques globales de la plage : plage de sable, plage de galets, plage en pente douce
nature_baignade	nature de la baignade : océanique, lacustre, bassin de baignade, bassin d'Arcachon
grand_territoire	Zone géographique sur laquelle se trouve la plage (Côte Basque, Nord Landes, Bassin d'Arcachon..)
source	Source de laquelle viennent les données (SIRTAQUI, Charente tourisme..)
date_ouverture	Date d'ouverture de la surveillance de la plage
date_fermeture	Date de fermeture de la surveillance de la plage
stationnement	Véhicules pour lesquels le stationnement est possible (ex : voitures, vélos, camping-car..)
date_detail	Détails sur les dates d'ouverture et de fermeture avec les horaires d'ouverture
equipements_services	Equipements disponibles sur la plage (restaurant)
id_plage_gip	Identifiant de la plage pour le GIP littoral
long	longitude
lat	latitude
insee_dep	Numéro insee du département où se situe la plage
iframe_parking	Plage équipée ou pas d'un capteur sur le parking
desc_acces_handi	Description de l'accessibilité pour les personnes à mobilité réduite
Plage surveillée	Plage surveillée ou non
descriptif	Description de la plage en quelques phrases

Pression touristique	Plage soumise à pression touristique ou pas
payant	Parking payant ou pas
influence	On peut avoir les données suivantes : Landes, BAB sud lan, Bx Bassin, Nord-Medoc
pression métropole	Plage soumise à la pression d'une métropole ou pas
risque incendie	Plage exposée ou pas à des risques incendie en période de forte sécheresse
accessibilité pmr	Plage accessible pour les personnes à mobilité réduite
accessibilité piéton	Plage accessible à pied ou non (distance maximale d'accessibilité piéton à définir)
accessibilité vélo	Niveau d'accessibilité à vélo (passable, bien, très bien)
accessibilité voiture	Plage accessible en voiture ou pas
capacité parking voiture	Nombre de places totales du ou des parking(s) voiture
capacité parking vélo	Nombre de places totales du ou des parking(s) vélo
impact des marées	Plage soumise à marée ou pas (3 niveaux : impact sur la surface de la plage disponible à marée haute / pas d'impact / pas de baignade possible à marée basse)
usage dominant	Activité dominante sur la plage
surface plage	Surface de la plage
plage équipée pour le comptage	Présence ou pas d'une caméra pour effectuer du comptage
urbanisation	Notion liée à la densité de bâtiments dans un rayon en arrière de la plage
présence de commerce	Présence de commerces permanents
nombre de parkings voiture	Nombre de parking(s) voiture
sports nautiques	Sports nautiques présents sur la plage
jeux pour enfants	Présence ou pas de jeux pour enfants sur la plage
équipement touristique	Liste des équipements touristiques présents sur la plage
nombre de lits	Nombre de lits dans une distance de 10 minutes à pied autour de la plage
population commune littorale	Population de la commune sur laquelle se situe la plage
proximité métropole	Proximité à la métropole la plus proche

Annexe 2 : Analyse du « silhouette score » pour 3 clusters



Annexe 3 : Liste des plages par clusters

Cluster 0 : essentiellement des plages océaniques, “naturelles”, relativement isolées et peu équipées.

Nom plage	Commune	Nom plage	Commune
Plage de Saint Nicolas	Le Verdon sur Mer	Plage Nord	Fouras-les-Bains
Plage Sud	Vendays Montalivet	Plage de La-Brée-les-Bains	La-Brée-les-Bains
Plage du Pin Sec	Naujac sur Mer	Plage du Peu Ragot	La Couarde-sur-Mer
Plage Nord	Lacanau	Plage de l'Arnérault	La Flotte
Plage du Truc Vert	Lege Cap Ferret	Plage de Chef de Baie	La Rochelle
Plage de Lespecier	Mimizan	Plage de la Pointe Espagnole	La Tremblade
Plage du Courant	Mimizan	Plage du Phare de la Courbe	La Tremblade
Plage de la Lette Blanche	Vielle Saint Girons	Plage de Gros Jonc	Le Bois-Plage-en-Ré
Plage des Chênes Lièges	Moliets et Maa	Plage des Gollandières	Le Bois-Plage-en-Ré
Plage Sud	Messanges	Plage de la Phibie	Le Château-D'Oléron
Plage des Casernes	Seignosse	Plage de la Giraudière	Le Grand-Village-Plage
Plage de la Gravière	Soorts Hossegor	Plage de la Garonne	Lège Cap Ferret
Plage du Pavillon Royal	Bidart	Plage de La Palmyre	Les Mathes
Plage du Fort de Socoa	Ciboure	Plage Les Vergnes	Meschers-sur-Gironde
Plage de la Platère	Angoulins	Plage Les Nonnes	Meschers-sur-Gironde
La plage	Châtelailon-Plage	Plage de Suzac	Meschers-sur-Gironde
Grande plage	Fouras-les-Bains	Plage des Anses	Port-des-Barques

Plage sud de Rivedoux-plage	Rivedoux-Plage	Plage de Plaisance	Saint-Georges-d'Oléron
Plage de la Conche des Baleines	Saint-Clément-des-Baleines	Plage des Sables Vignier	Saint-Georges-d'Oléron
Plage des Huttes	Saint-Denis-d'Oléron	Plage de Chaucre	Saint-Georges-d'Oléron
Plage de la Boirie	Saint-Denis-d'Oléron	Grande Plage Nord	Saint-Georges-de-Didonne
Plage de Boyardville	Saint-Georges-d'Oléron	Plage de la Cible	Saint-Martin-en-Ré
Plage de Matha	Saint-Pierre-d'Oléron	Plage de la Grande Côte	Saint-Palais-sur-Mer
Plage de Gatseau	Saint-Trojan-les-Bains	Plage de Montamer	Sainte-Marie-de-Ré
La Grande Plage	Saint-Trojan-les-Bains	Plage Nord	Soulac sur Mer
Plage de la Gautrelle	Saint-Georges-d'Oléron	Plage des Pins de Cordouan	Les Mathes
Plage des Bonnes	Saint-Georges-d'Oléron	Plage du Bureau	Saint-Palais-sur-Mer
Plage la Bouverie	La Tremblade	Plage Les Combots	Saint-Palais-sur-Mer
Plage Sud	Fouras-les-Bains	Plage du Petit Bec	Les Portes-en-Ré
Plage de l'Anse de la Croix	Ile-d'Aix	Grande Plage Sud	Saint-Georges-de-Didonne
Plage de la Rémigeasse	Dolus-d'Oléron	Plage des Seulières	Saint-Denis-D'Oléron
Plage de Vertbois	Dolus-d'Oléron		

Cluster 1 : essentiellement des plages océaniques, bien équipées.

Nom plage	Commune	Nom plage	Commune
Plage de la Chambrette	Le Verdon sur Mer	Plage Centrale	Labenne
Plage Centrale	Soulac sur Mer	Plage Centrale	Ondres

Plage Sud	Soulac sur Mer	Plage du Métro	Tarnos
Plage de l'Amélie	Soulac sur Mer	Plage de la Digue	Tarnos
Plage Centrale	Hourtin	Plage de La Barre	Anglet
Plage Centrale	Carcans	Plage des Cavaliers	Anglet
Plage Centrale	Lacanau	Plage de l'Océan	Anglet
Plage Sud	Lacanau	Plage de la Madrague	Anglet
Plage Super Sud	Lacanau	Plage des Corsaires	Anglet
Plage du Gressier	Le Porge	Plage de Marinella	Anglet
Plage du Grand Crohot	Lege Cap Ferret	Plage de la Petit Chambre d'Amour	Anglet
Plage de l'Horizon	Lege Cap Ferret	Plage des Sables d'Or	Anglet
Plage Pereire	Arcachon	Plage Marbella	Biarritz
Plage d'Arcachon	Arcachon	Plage du Miramar	Biarritz
Plage du Moulleau	Arcachon	Grande Plage	Biarritz
Plage de la Corniche	La Teste de Buch	Plage de la Milady	Biarritz
Plage du Petit Nice	La Teste de Buch	Plage de la Côte des Basques	Biarritz
Plage de la Lagune	La Teste de Buch	Plage du Port Vieux	Biarritz
Plage de la Salie-Nord	La Teste de Buch	Plage d'Ilbarritz	Bidart
Plage du Vivier	Biscarrosse	Plage de Parlementia	Bidart
Plage Nord	Biscarrosse	Plage d'Erretegia	Bidart
Plage Centrale	Biscarrosse	Plage du Centre	Bidart
Plage Sud	Biscarrosse	Plage de l'Uhabia	Bidart
Plage Sud	Mimizan	Plage de Mayarco	Saint Jean de luz
Plage Centrale	Moliets et Maa	Plage de Lafitenia	Saint Jean de luz
Plage Nord	Mimizan	Plage d'Erromardie	Saint Jean de luz
Plage Centrale	Vendays Montalivet	Grande plage Pergola - Digue aux chevaux	Saint Jean de luz
Plage de Contis	Saint Julien en Born	Plage Sud	Saint Jean de luz

Plage du Cap de l'Homy	Lit Et Mixe	Plage de Socoa	Ciboure
Saint Girons Plage	Vielle Saint Girons	Plage des 2 jumeaux	Hendaye
Plage du Col Vert	Vielle Saint Girons	Plage pointe de Sokoburu	Hendaye
Plage Centrale	Messanges	Plage des Minimés	La Rochelle
Plage Nord des Sablières	Vieux Boucau	Plage de la Concurrence	La Rochelle
Plage Centrale	Vieux Boucau	Plage du Pigeonnier	Royan
Plage de l'Océan	Soustons	Plage du Chay	Royan
Plage du Gurp	Grayan et L'Hopital	Plage de la Grande Conche	Royan
Plage du Penon	Seignosse	Plage de Foncillon	Royan
Plage des Bourdaines	Seignosse	Plage de Pontailac	Royan
Plage des Estagnots	Seignosse	Plage de Vallières	Saint-Georges-de-Didonne
Plage Centrale	Soorts Hossegor	Plage de Nauzan	Vaux-sur-Mer
Plage Notre Dame	Capbreton	Plage Sud	Soorts Hossegor
Plage de la Piste	Capbreton	Plage Naturiste	Soorts Hossegor
Plage des Océanides	Capbreton	Plage du Prévent	Capbreton
Plage Centrale	Capbreton	Plage de l'Estacade	Capbreton
Plage Santocha	Capbreton	Plage de Senix	Saint Jean de Luz

Cluster 2 : essentiellement des plages lacustres et bassins de baignade.

Nom plage	Commune	Nom plage	Commune
Plage de Piqueyrot	Hourtin	Plage du Ponton	Aureilhan
Plage du Port	Hourtin	Plage du Lac	Leon
Plage du Montaut	Carcans	Plage du Lac Marin	Vieux Boucau
Plage de Bombannes	Carcans	Plage du Lac Marin La Sauvagine	Soustons
Plage de Maubuisson	Carcans	Plage du Rey	Soorts Hossegor

Plage de la Grande Escoure	Lacanau	Plage Blanche	Soorts Hossegor
Plage du Moutchic	Lacanau	Plage des Chênes Lièges	Soorts Hossegor
Plage de Saint Brice	Ares	Plage du Parc	Soorts Hossegor
Bassin de baignade de Lanton	Lanton	Plage du Lac	Saint Pée sur Nivelle
Bassin de baignade Emile Ortel	Audenge	Plage Laouga	La Teste de Buch
Bassin de baignade du Teich	Le Teich	Plage Mouteou Pipiou	Parentis en Born
Plage de La Hume	Gujan Mestras	Plage de Navarrosse	Biscarrosse
Plage du lac de Cazaux	La Teste de Buch	Plage du Port de Maguide	Biscarrosse
Plage de Caton	Sanguinet	Plage du Port de Gastes	Gastes
Plage du Pavillon	Sanguinet	Plage de Sainte Eulalie en Born	Sainte Eulalie en Born

Annexe 4 : Formules de calcul

- Calcul du coefficient du « silhouette score » :

- Un coefficient de cohésion $a(i)$ est calculé pour chaque point, c'est la distance moyenne entre le point i et l'ensemble des autres points de son cluster :

$$a(i) = \frac{1}{|I_k|-1} \sum_{j \in I_k, j \neq i} d(x^i, x^j)$$

avec $I_k = \{i \in [1, N] | C(i) = k\}$ l'ensemble des points du cluster k .

- Un coefficient de séparation $b(i)$ est calculé pour chaque point, c'est le minimum des distances moyennes du point i à l'ensemble des points d'un autre cluster :

$$b(i) = \min_{k' \neq k} \frac{1}{|I_{k'}|} \sum_{i' \in I_{k'}} d(x^i, x^{i'})$$

- Le coefficient silhouette $s_{sil}(i)$ du point i peut alors être calculé :

$$s_{sil}(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Si ce coefficient est négatif alors $a(i) > b(i)$: on est plus proche d'un autre cluster que de celui auquel il appartient : il est donc mal classé. Plus le point est proche des points de son cluster et plus il est séparé de ceux des autres clusters, plus le coefficient de silhouette sera grand.

- Le silhouette score S_{sil} peut finalement être déterminé, c'est la moyenne des moyennes des coefficients de silhouette au sein de chaque cluster :

$$S_{sil} = \frac{1}{K} \sum_{k=1}^K \frac{1}{|I_k|} \sum_{i \in I_k} s_{sil}(i), \text{ où } K \text{ est le nombre de clusters.}$$

- Formule de normalisation de données :

$$x_{norm,i} = \frac{x_i - \min_{j \in \{1, \dots, n\}}(x_j)}{\max_{j \in \{1, \dots, n\}}(x_j) - \min_{j \in \{1, \dots, n\}}(x_j)}, \text{ avec } n \text{ le nombre d'observations et } i \in \{1, \dots, n\}$$

GIP LITTORAL

2030



MEMBRES DU GIP

Services de l'Etat en région, conseil régional Nouvelle-Aquitaine, conseil départemental de la Charente-Maritime, conseil départemental de la Gironde, conseil départemental des Landes, conseil départemental des Pyrénées-Atlantiques, communauté d'agglomération de La Rochelle, communauté d'agglomération Rochefort Océan, communauté d'agglomération Royan Atlantique, communauté d'agglomération du Bassin d'Arcachon Nord, communauté d'agglomération Bassin d'Arcachon Sud, communauté d'agglomération Pays Basque, communauté de communes de l'île de Ré, communauté de communes de l'île d'Oléron, communauté de communes Bassin de Marennes, communauté de communes Médoc Atlantique, communauté de communes de la Médullienne, communauté de communes des Grands Lacs, communauté de communes de Mimizan, communauté de communes Côte Landes Nature, communauté de communes Marenne Adour Côte Sud, communauté de communes du Seignanx

GIP LITTORAL

11, avenue Pierre Mendès-France 33700 MÉRIGNAC

T. + 33 (0)5 56 13 26 28

contact@gip littoral.fr

www.giplittoral.fr

