

Prédire l'affluence de la fréquentation des sentiers à partir de l'historique des flux et des prévisions météo - cas d'étude du Parc national des Écrins

Prevedere l'affluenza sui sentieri sulla base dei dati storici relativi ai flussi di visitatori e delle previsioni meteorologiche - caso di studio del Parco Nazionale degli Écrins

par Julien Rohaut et Marc Langenbach (Université Grenoble-Alpes)



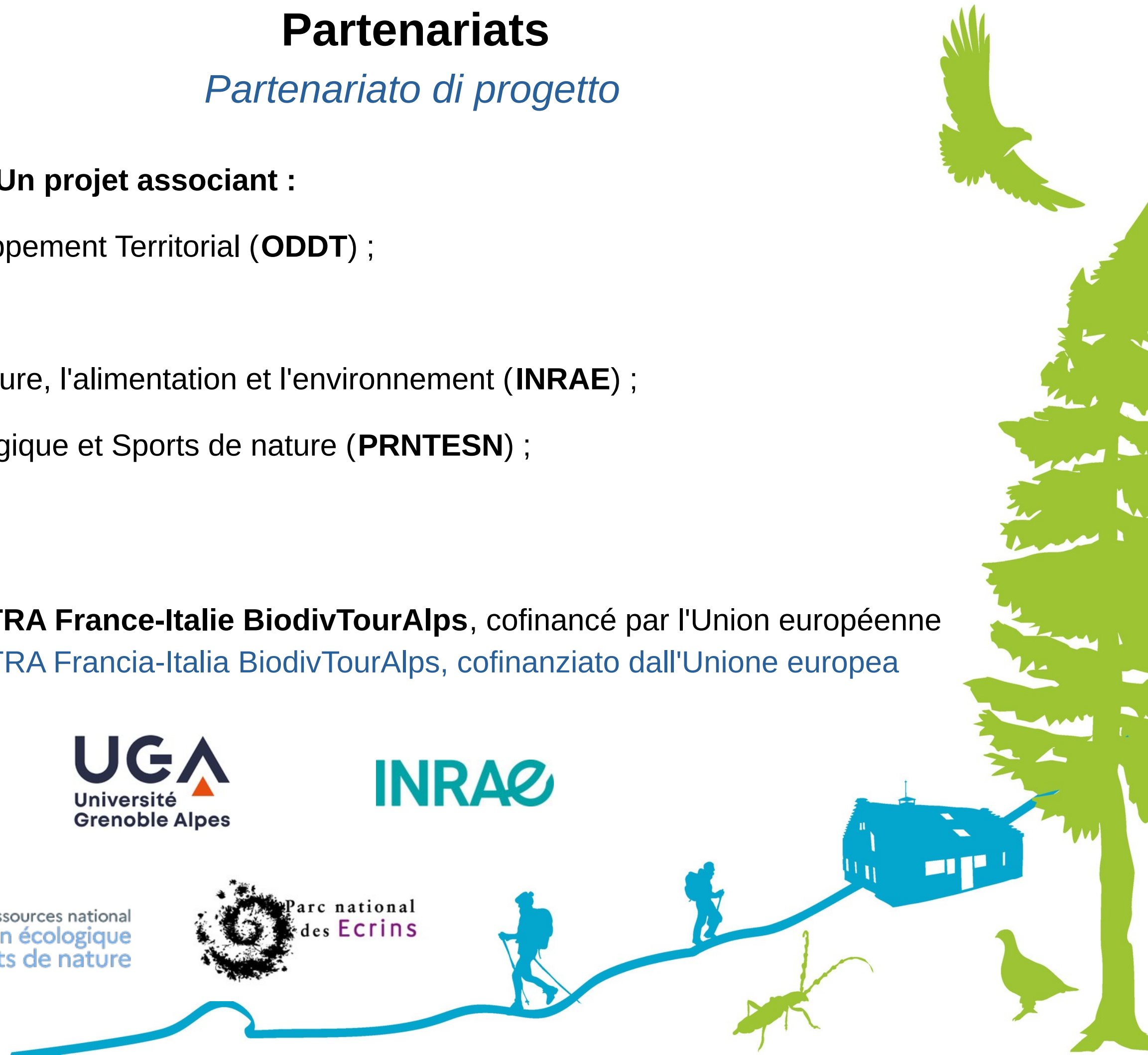
Partenariats

Partenariato di progetto

Un projet associant :

- l'Observation des Dynamiques et du Développement Territorial (**ODDT**) ;
- l'Université Grenoble-Alpes (**UGA**) ;
- L'Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (**INRAE**) ;
- Le Pôle ressources national Transition écologique et Sports de nature (**PRNTESN**) ;
- le Parc national des Écrins.

→ Réalisé dans le cadre du programme **ALCOTRA France-Italie BiodivTourAlps**, cofinancé par l'Union européenne
Realizzato nell'ambito del programma ALCOTRA Francia-Italia BiodivTourAlps, cofinanziato dall'Unione europea



Pourquoi prédire la fréquentation ?

Perché prevedere l'affluenza?

- Prédire la fréquentation sur des sentiers permettrait de comprendre et d'anticiper le comportement des pratiquant·e·s de sports de nature / Prevedere l'affluenza sui sentieri consentirebbe di comprendere e anticipare il comportamento di chi pratica sport all'aria aperta;
- Des randonneur·euse·s pourraient choisir de se reporter vers des itinéraires moins fréquentés si l'information leur était accessible en amont / Gli escursionisti potrebbero scegliere di optare per percorsi meno frequentati se potessero disporre di tali informazioni in anticipo;
- Anticiper les flux permettrait de mieux protéger les zones naturelles sensibles à forte pression anthropique / Prevedere i flussi consentirebbe di proteggere meglio le aree naturali sensibili soggette a forte pressione antropica;
- Les traces géonumériques (Strava, OutdoorVision, etc.) et les écompteurs offrent une nouvelle solution, encore sous-exploitée, pour anticiper et gérer ces flux (Rohaut et al., 2025) / Le tracce geolocalizzate (Strava, OutdoorVision, ecc.) e i contatori ecologici offrono una nuova soluzione, ancora poco sfruttata, per prevedere e gestire questi flussi (Rohaut et al., 2025).

Gestionnaires

Comprendre, anticiper et gérer les flux de visiteurs (prévention, aménagements...)

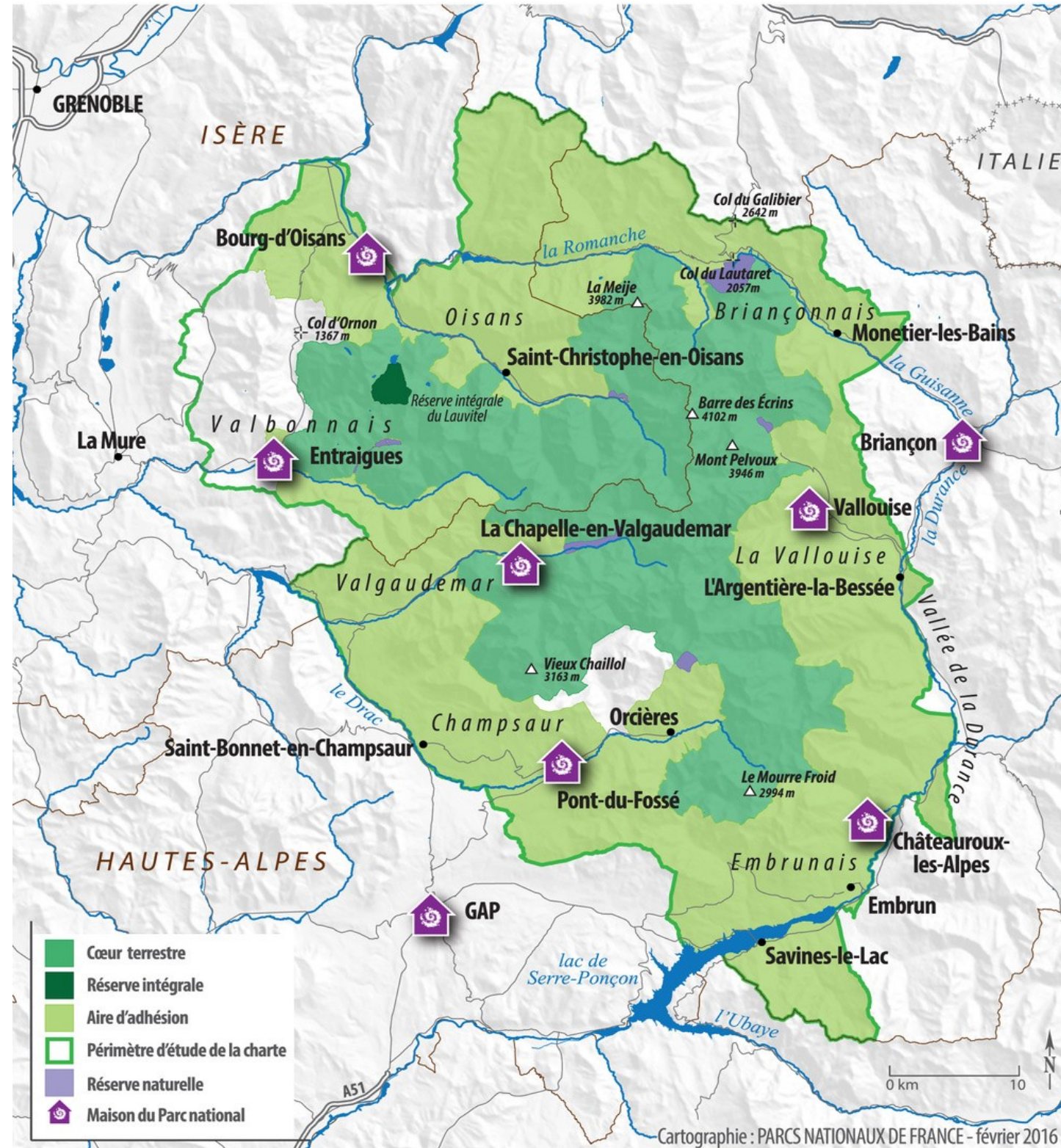


Grand Public

Informar de l'intensité de fréquentation à un endroit, un jour donné et à une heure donnée



Le Parc national des Écrins



En quelques chiffres

- 93 000 ha en cœur de parc
- 160 600 ha en aire d'adhésion
- 700 km de sentiers entretenus
- 274 oiseaux et 94 mammifères recensés
- 2 314 espèces végétales

PNE, Rapport d'activité 2025



→ 700 km de sentiers, notamment au sein de milieux naturels sensibles → un territoire où un outil de prédiction de la fréquentation constituerait une aide à la décision pour les gestionnaires

Données de fréquentation : traces et écocompteurs

Dati di frequentazione: tracciati e contatori ecologici

- **STRAVA METRO** : Plateforme de Strava qui anonymise et agrège les données d'activité de ses utilisateur·rice·s selon deux grands types (**Pédestre** : Run/Walk/Hike ; **Vélo** : Bike/E-Bike). Les données sont agrégées sur les tronçons d'OpenStreetMap (OSM) ; → des pratiquant·e·s au **profil plutôt sportif·ve·s**
- **OUTDOORVISION** : Plateforme à but non lucratif portée par le **Pôle ressources national Transition écologique et Sports de nature**. Elle agrège les traces GPS des pratiquant·e·s de sports de nature (marche, course, vélo et ski) pour aider collectivités et gestionnaires d'espaces naturels à mieux comprendre les usages sur le terrain ; → des **pratiquant·e·s de profil "grand public"**
- **ECOCOMPTEURS** : Capteurs physiques déployés sur les sentiers du PNE comptabilisant le nombre de passages.

	OutdoorVision	Strava Metro	Écocompteurs
Type	Traces individuelles	Données agrégées sur segments de tronçons OSM	Comptage terrain
Période	2020-2024	2021-2025	2020-2024
Volumétrie	24 905 traces	14 392 km de réseau	26 compteurs
Activité utilisée	Marche	Marche/course/randonnée	Piétons sur sentiers
Limite principale	Quantitative	Biais population + boîte noire algorithmique	Ponctuel et localisé



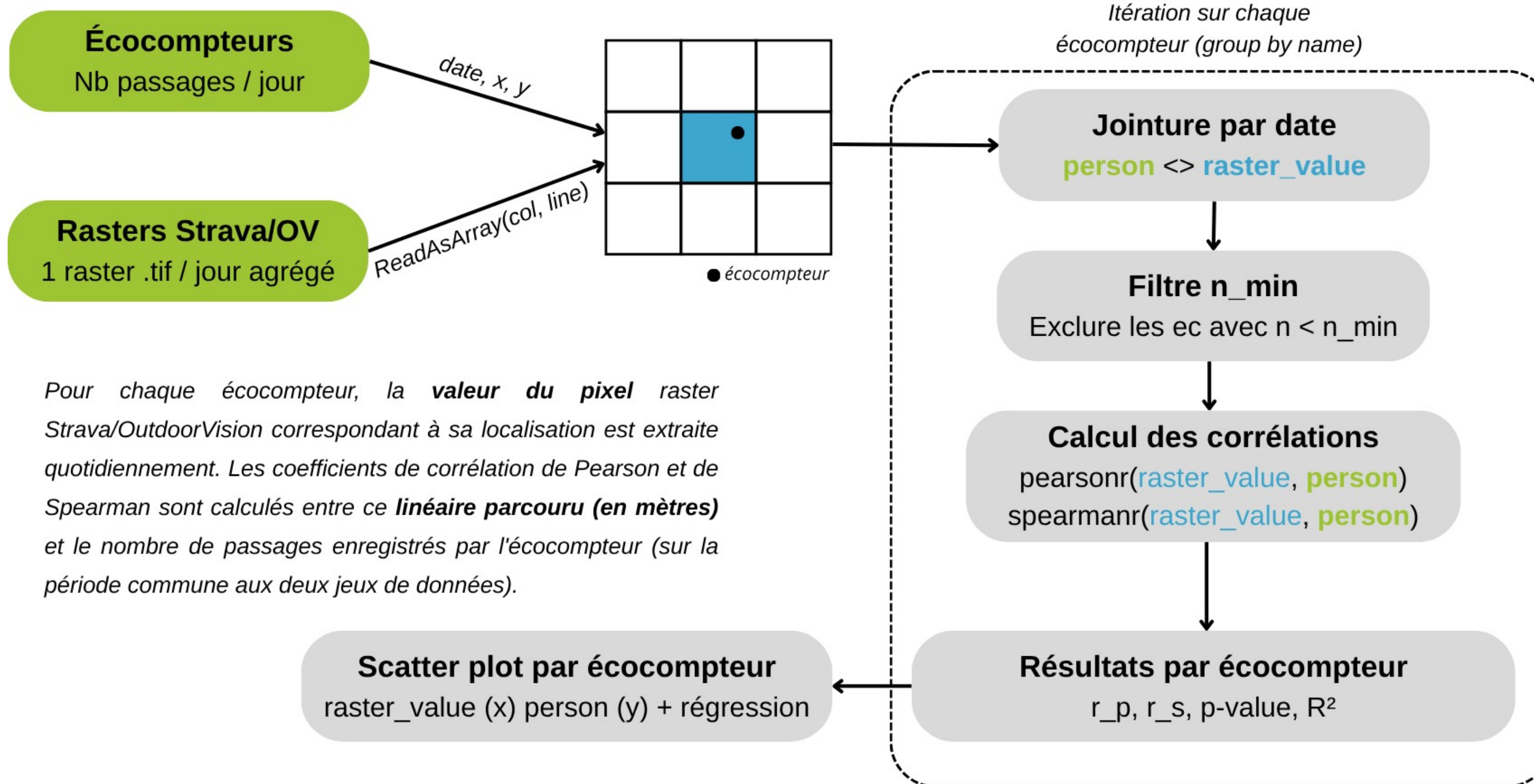
Représentativité des traces dans la fréquentation

Rappresentatività delle tracce nella frequentazione

Peut-on réellement “faire confiance” aux traces géonumériques pour estimer une fréquentation réelle ?

Ci si può davvero «fidare» dei dati geodigitali per stimare l'affluenza effettiva?

Méthode pour mesurer les corrélations entre données issues des traces et données écompteurs



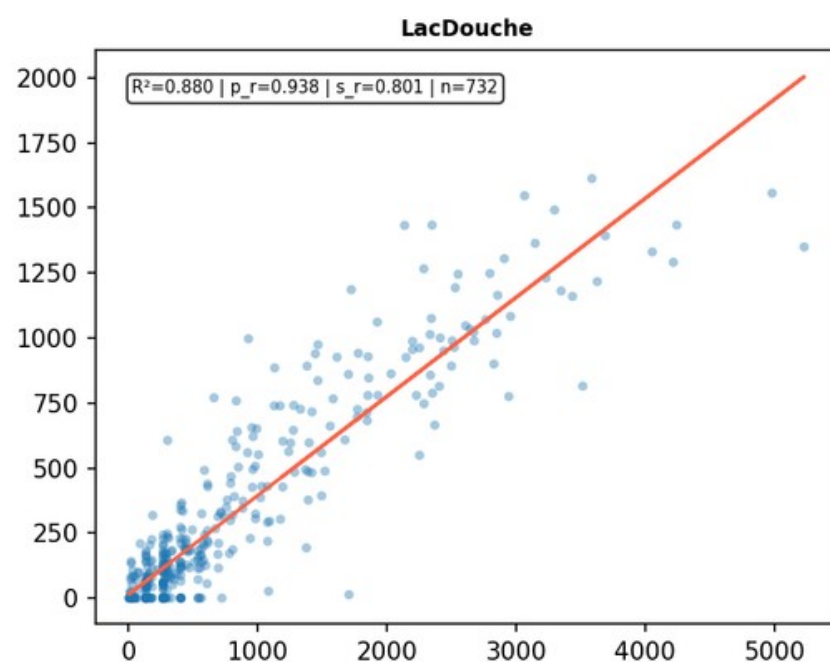
Représentativité des traces dans la fréquentation

Rappresentatività delle tracce nella frequentazione

Groupe 1

Relation linéaire forte
($r > 0.7$)

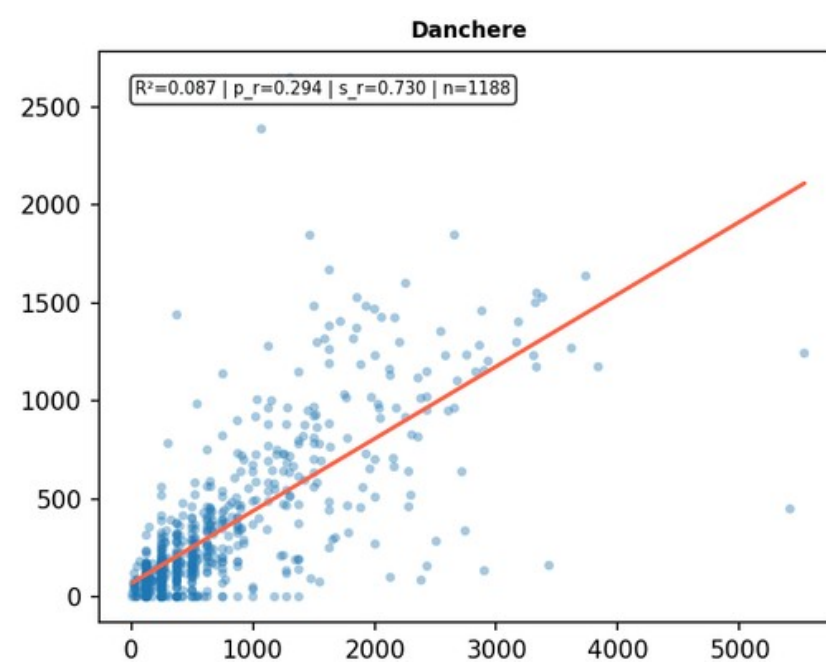
→ **13 écompteurs**
pearson r moy = 0.860
spearman r moy = 0.773



Groupe 2

Relation modérée voire non linéaire
($0.5 < r < 0.7$)

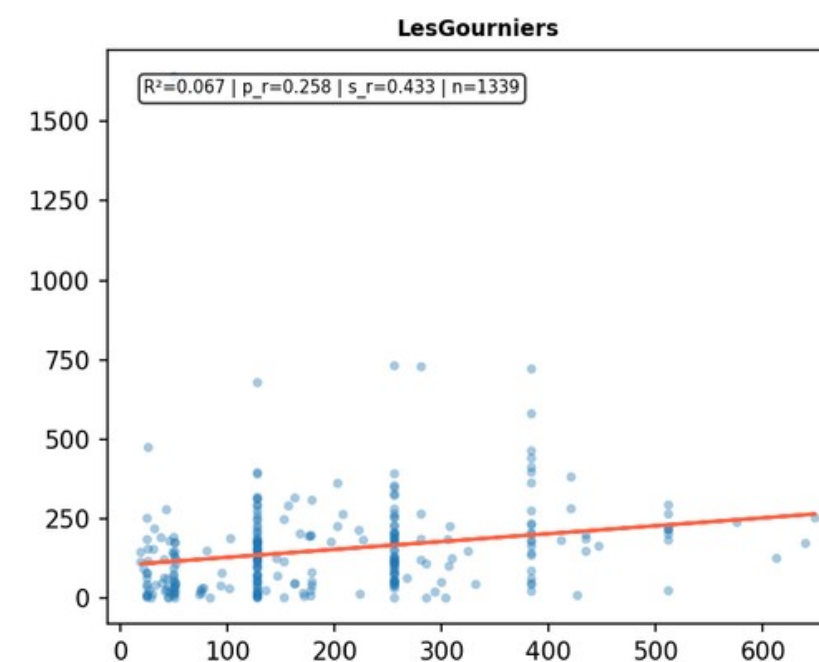
→ **5 écompteurs**
pearson r moy = 0.592
spearman r moy = 0.619



Groupe 3

Relation faible voire inexistante
($r < 0.5$)

→ **7 écompteurs**
pearson r moy = 0.198
spearman r moy = 0.283



Classement basé sur le coefficient de Pearson r (Strava/OutdoorVision, période entière)

- Les traces agrégées Strava/OutdoorVision constituent un proxy fiable de la fréquentation réelle sur la majorité des sentiers du PNE couverts par au moins un écompteur
- Le tracce aggregate di Strava/OutdoorVision costituiscono un indicatore affidabile dell'affluenza effettiva sulla maggior parte dei sentieri del PNE coperti da almeno un contatore ecologico

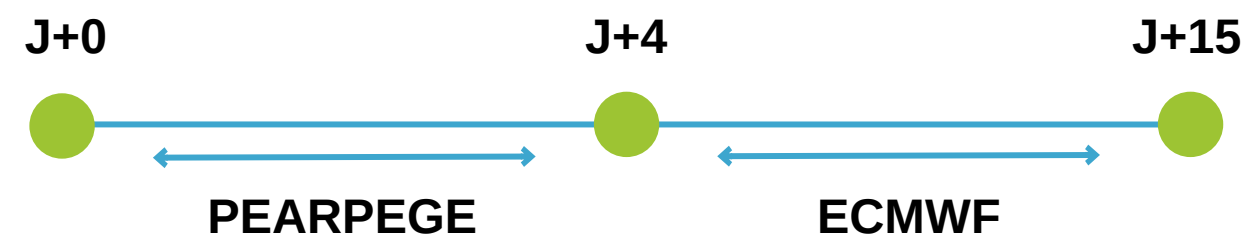
Les données météo : historique et prévisionnel

Dati meteorologici: dati storici e previsioni

	SAFRAN	PEARPEGE	ECMWF
Usage	Entraînement du modèle	Prédiction	Prédiction
Pas de temps	Réanalyse depuis 1958	J+1 à J+4	J+5 à J+15
Résolution spatiale	8x8 km	11x11 km	28 km
Approche	Réanalyse déterministe	Ensemble (médiane <i>n</i> membres)	Ensemble (médiane <i>n</i> membres)
Source	API Safran-Géos (Météo-France)	API PEARPEGE (Météo-France)	Module Python ecmwf.opendata (ECMWF*)

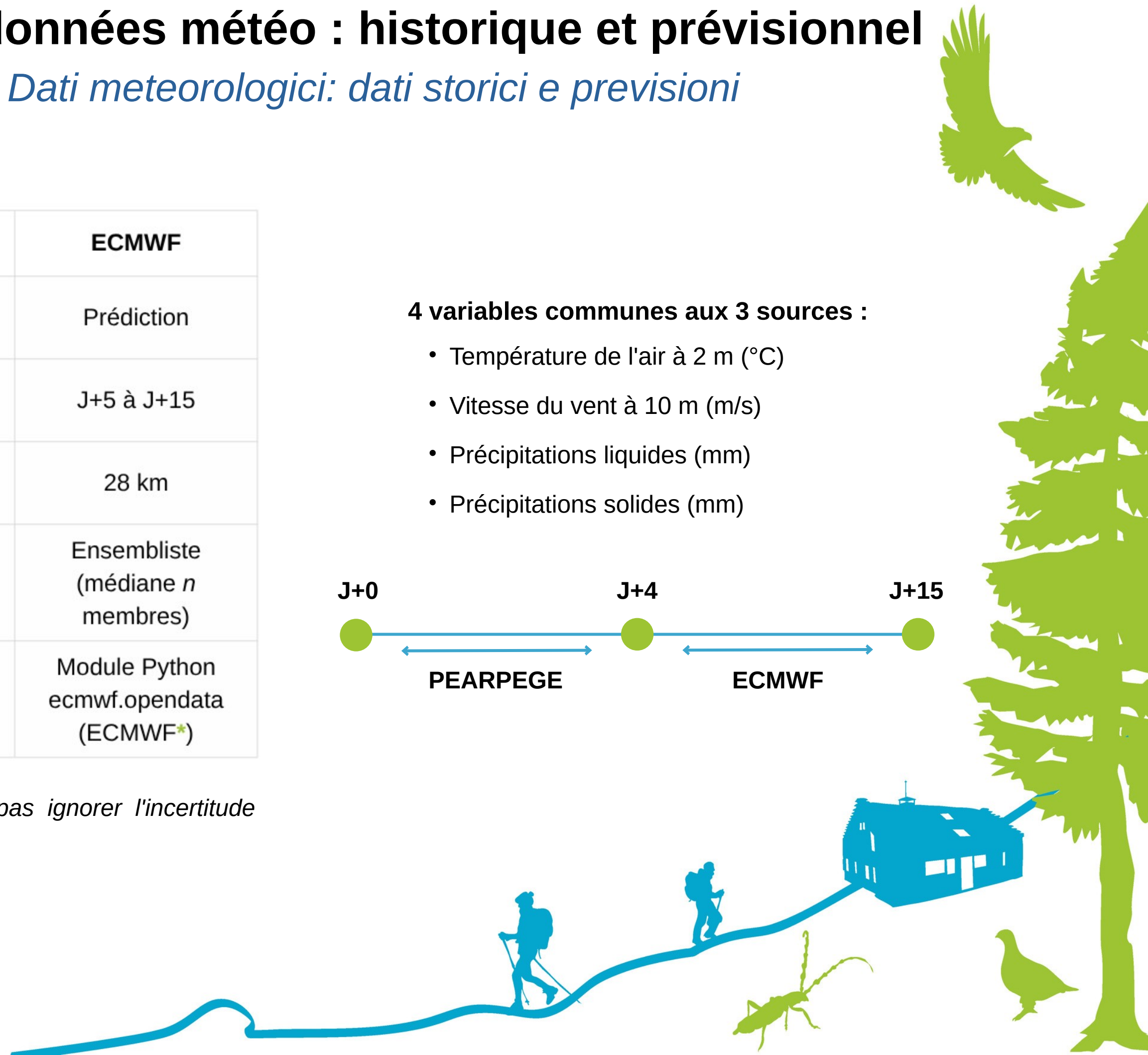
4 variables communes aux 3 sources :

- Température de l'air à 2 m (°C)
- Vitesse du vent à 10 m (m/s)
- Précipitations liquides (mm)
- Précipitations solides (mm)



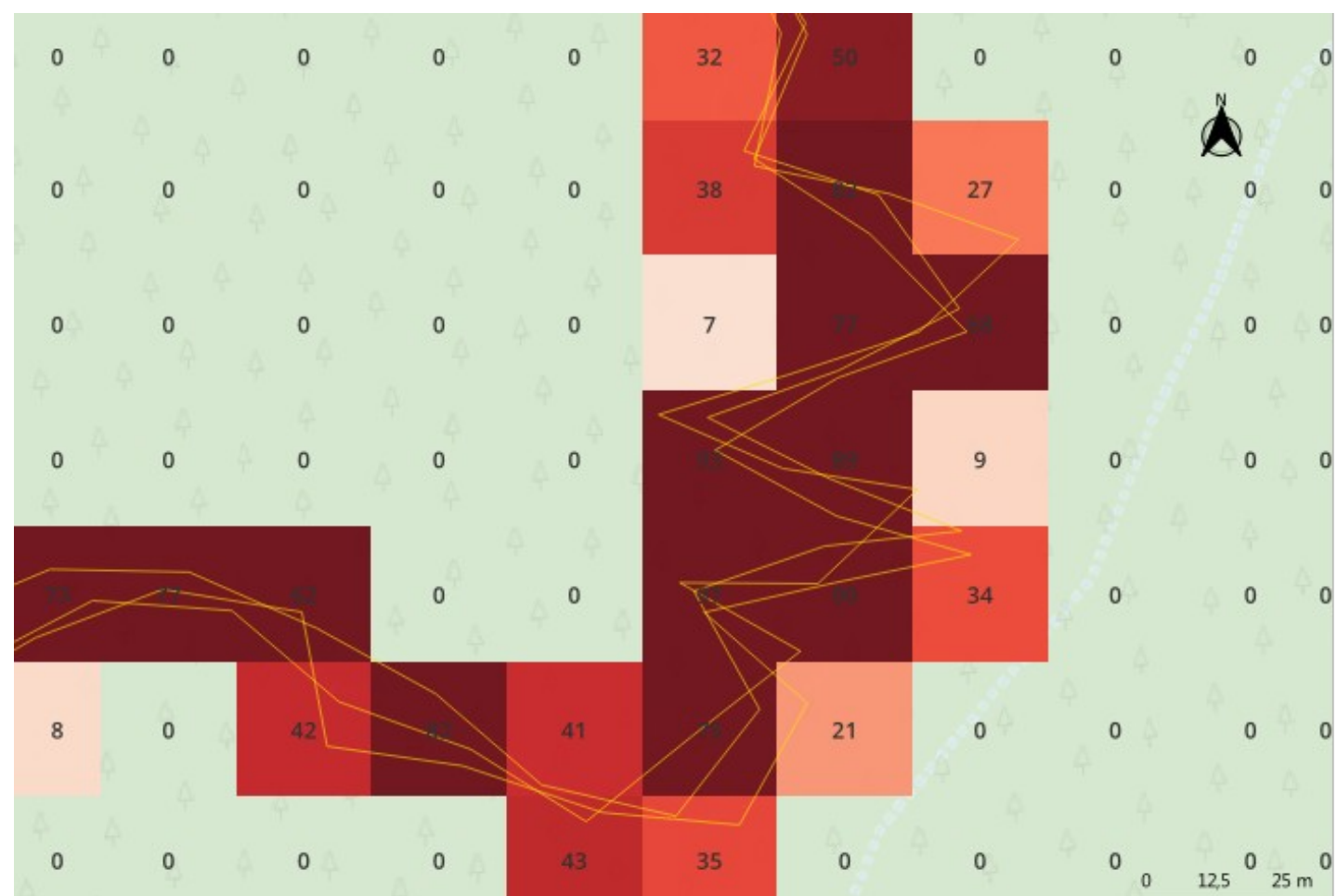
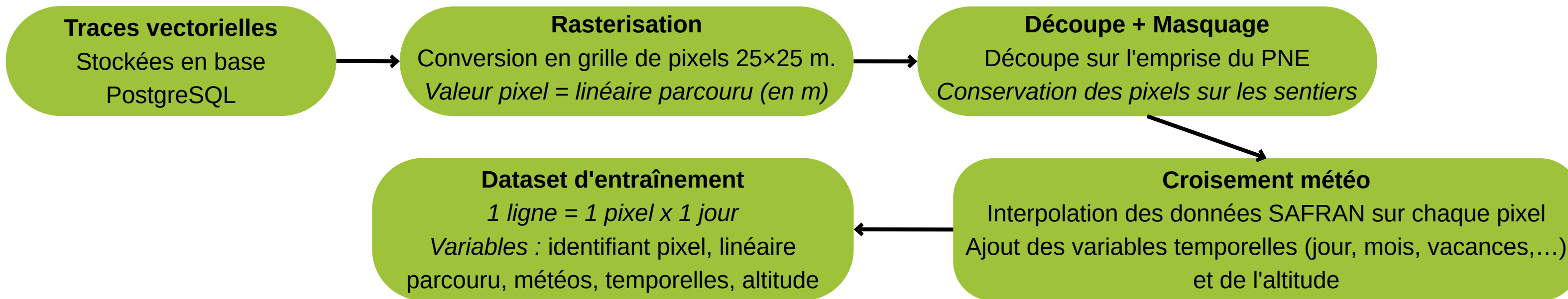
L'approche ensembliste (PEARPEGE, ECMWF) permet de ne pas ignorer l'incertitude météo

*European Centre for Medium-Range Weather Forecasts



Chaîne de traitement : rasterisation - pipeline

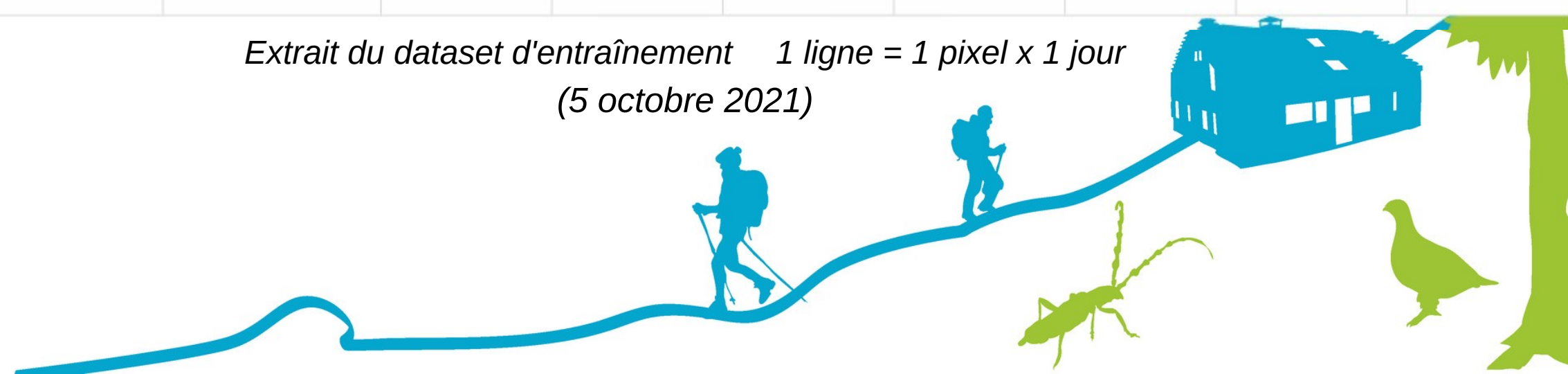
Catena di elaborazione: rasterizzazione - pipeline



Exemple d'un extrait de traces rasterisé du jeu de données OutdoorVision (12 juillet 2020), avec la valeur du linéaire parcouru de chaque pixel de 25x25 (en mètres)

pixel_id	date	freq (m)	t_q (°C)	ff_q (m/s)	preliq_q (mm)	prenei_q (mm)	weekend	weekday
101	05/10/2021	50	1,7	3,5	1,7	2	0	2
1 025	05/10/2021	73	1,7	3,5	1,7	2	0	2
2 003	05/10/2021	81	1,7	3,5	1,7	2	0	2

Extrait du dataset d'entraînement 1 ligne = 1 pixel x 1 jour (5 octobre 2021)



Le modèle prédictif XGBoost - entraînement

Il modello predittivo XGBoost - addestramento

XGBoost → algorithme d'apprentissage automatique par arbres de décision

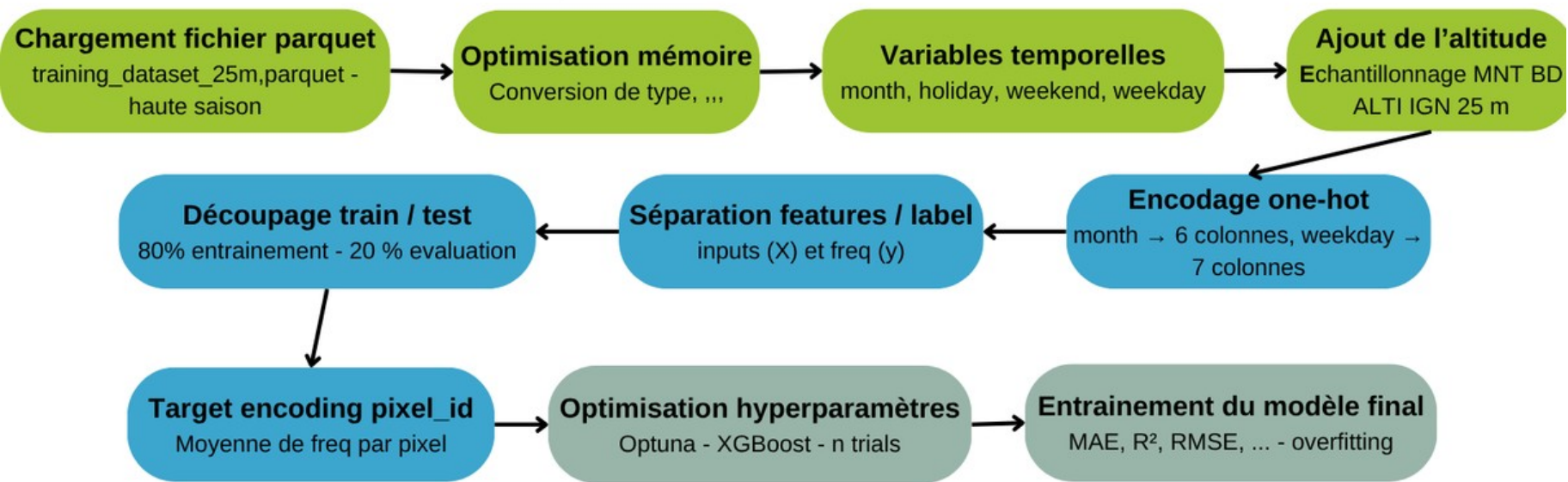
Entraîné sur 4 ans de données (2021-2024), haute saison (mai-octobre)

Étapes résumées de la construction du modèle : Données en entrée → Préparation → Optimisation automatique → Modèle retenu

Variables d'entrée (inputs) :

- 12 combinaisons de variables d'entrée testées
 - > 500 entraînements au total
 - environ 1 300 h de calcul cumulé
 - 75 959 pixels prédits par jour
- Localisation du pixel
 - Altitude
 - Température
 - Vitesse du vent
 - Précipitations liquides et solides
 - Mois, jour de semaine, weekend, jour férié

Étapes de construction du modèle

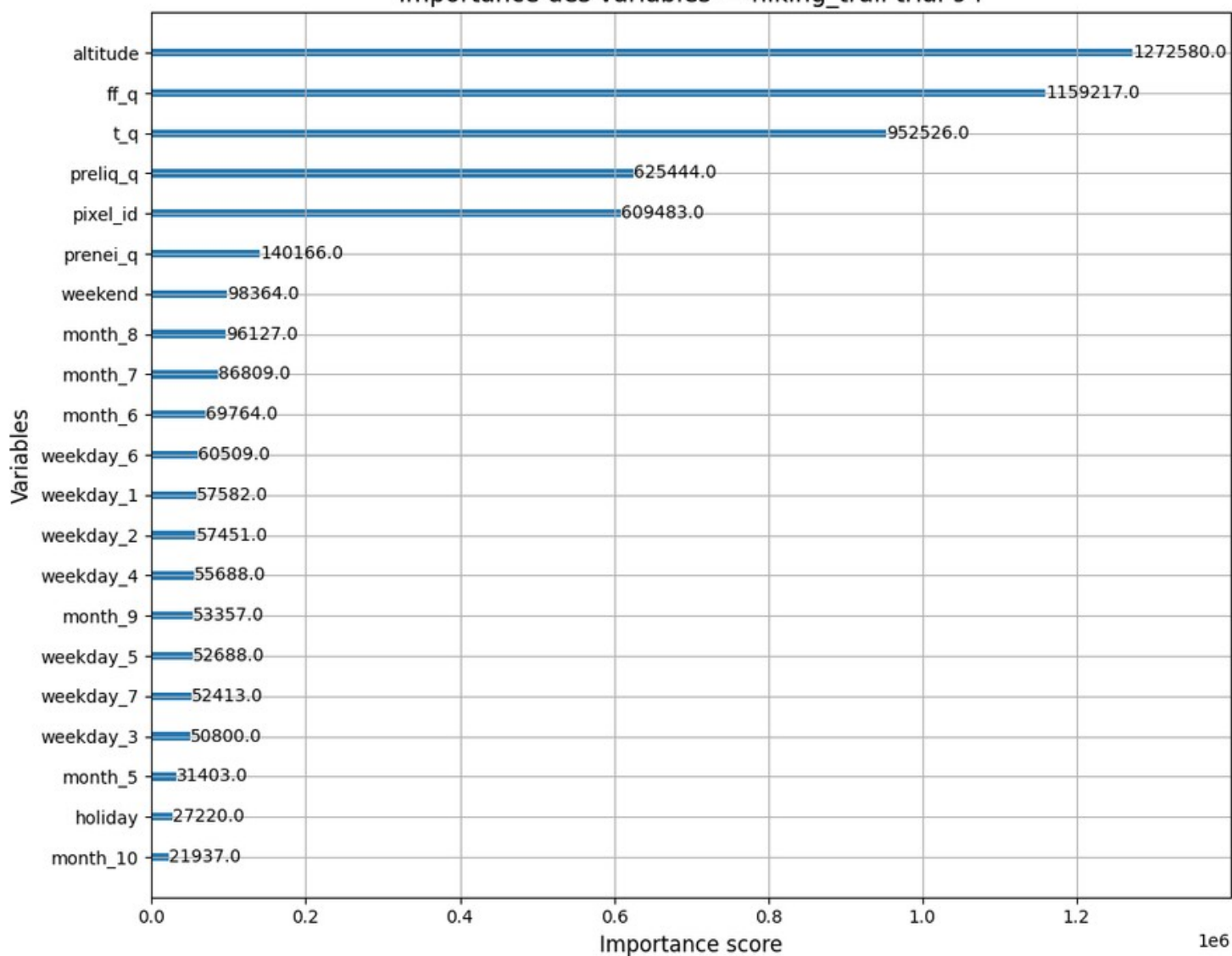


Le modèle prédictif XGBoost - résultats

Il modello predittivo XGBoost - risultati

Un modèle est évalué sur sa capacité à prédire des données “qu'il n'a jamais vues” (jeu de test, 20 % des données). Parmi les indicateurs clés : le R^2 mesure la part de variabilité expliquée (plus proche de 1 = meilleur) ; la MAE mesure l'erreur moyenne de prédiction en mètres (plus proche de 0 = meilleur).

Importance des variables — hiking_trail trial 94

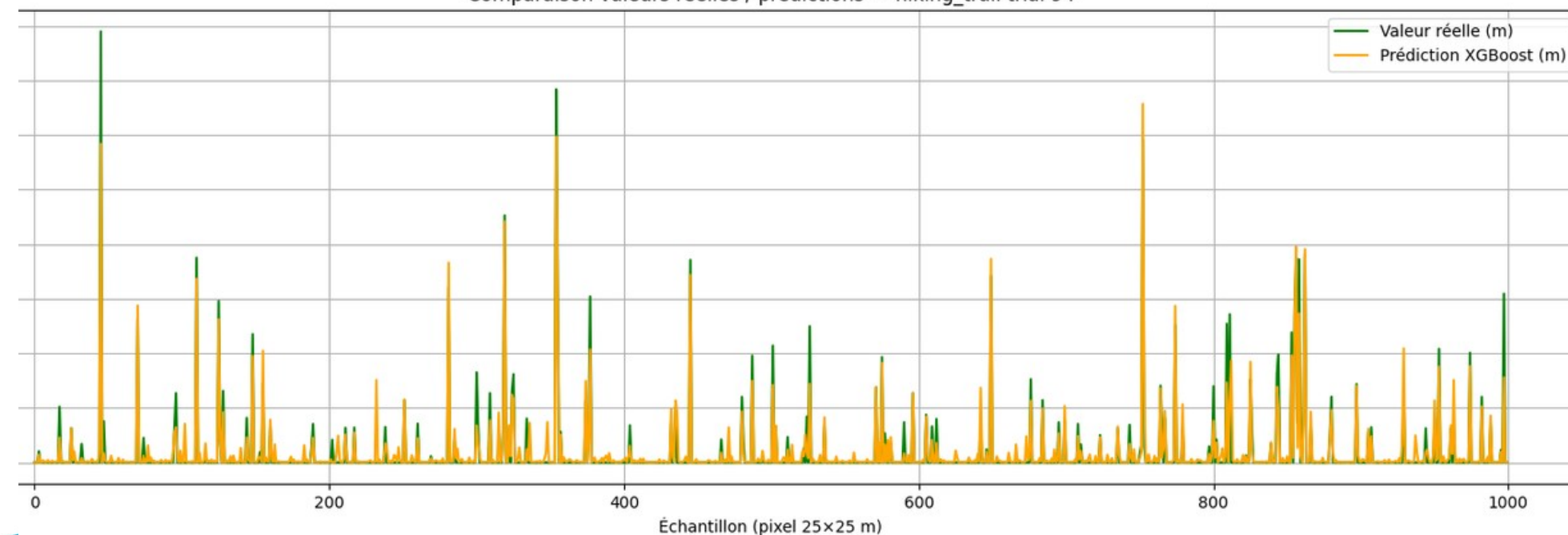


Le score d'importance mesure combien chaque variable a contribué à réduire l'erreur de prédiction à travers l'ensemble des arbres du modèle

La MAPE>0 (Erreur relative en % sur les pixels fréquentés) est tirée vers le haut par les pics de fréquentation des valeurs extrêmes rares mais difficiles à prédire

Métrique	Test	Interprétation
R^2	0,847	85 % de la variabilité expliquée
MAE	18,2 m	Erreur moyenne par pixel
MAPE>0	52,90%	Erreur relative sur les pixels fréquentés

Comparaison valeurs réelles / prédictions — hiking_trail trial 94



Les seuils de fréquentation

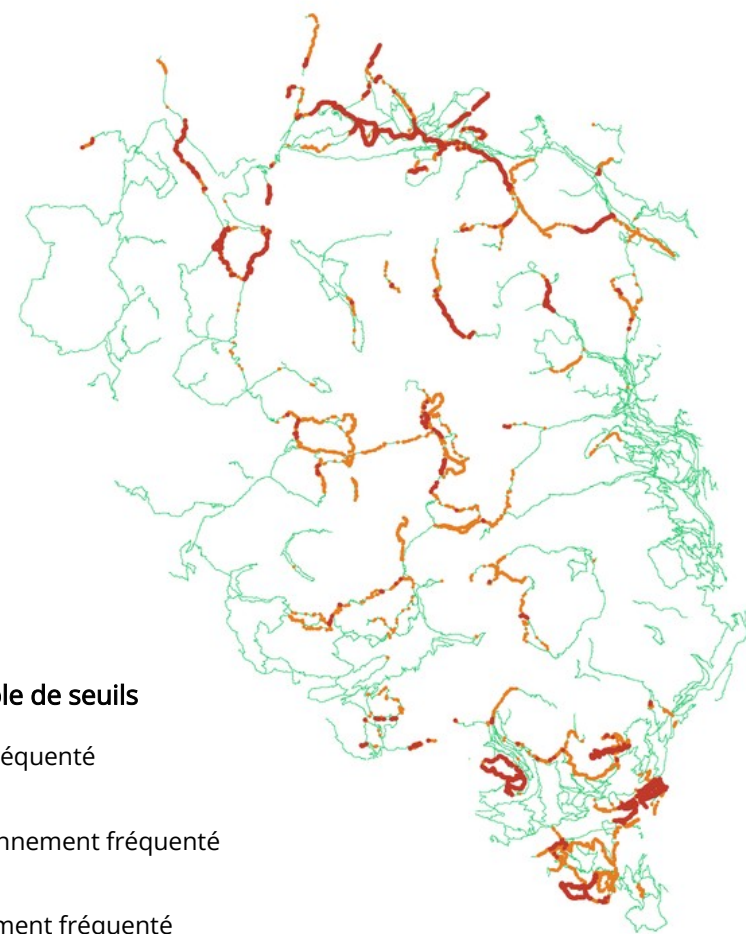
Le soglie di frequentazione

Le modèle prédit un linéaire parcouru en mètres par pixel → c'est une valeur qui ne parle pas au grand public. Des seuils permettent de traduire cette valeur en niveau de fréquentation

Pour donner une information indicative de la fréquentation à J+15, la méthode suivante est en cours de réalisation :

- Découpage du parc en zones homogènes selon la localisation et la fréquentation des sentiers ;
- Calcul de seuils propres à chaque zone - un sentier très fréquenté et un sentier isolé n'ont pas la même échelle de référence ;
- Classification de la valeur prédite en 3 niveaux (feu tricolore) : Peu fréquenté / Moyennement fréquenté / Fortement fréquenté

Exemple de prédiction (dimanche 14 juin)

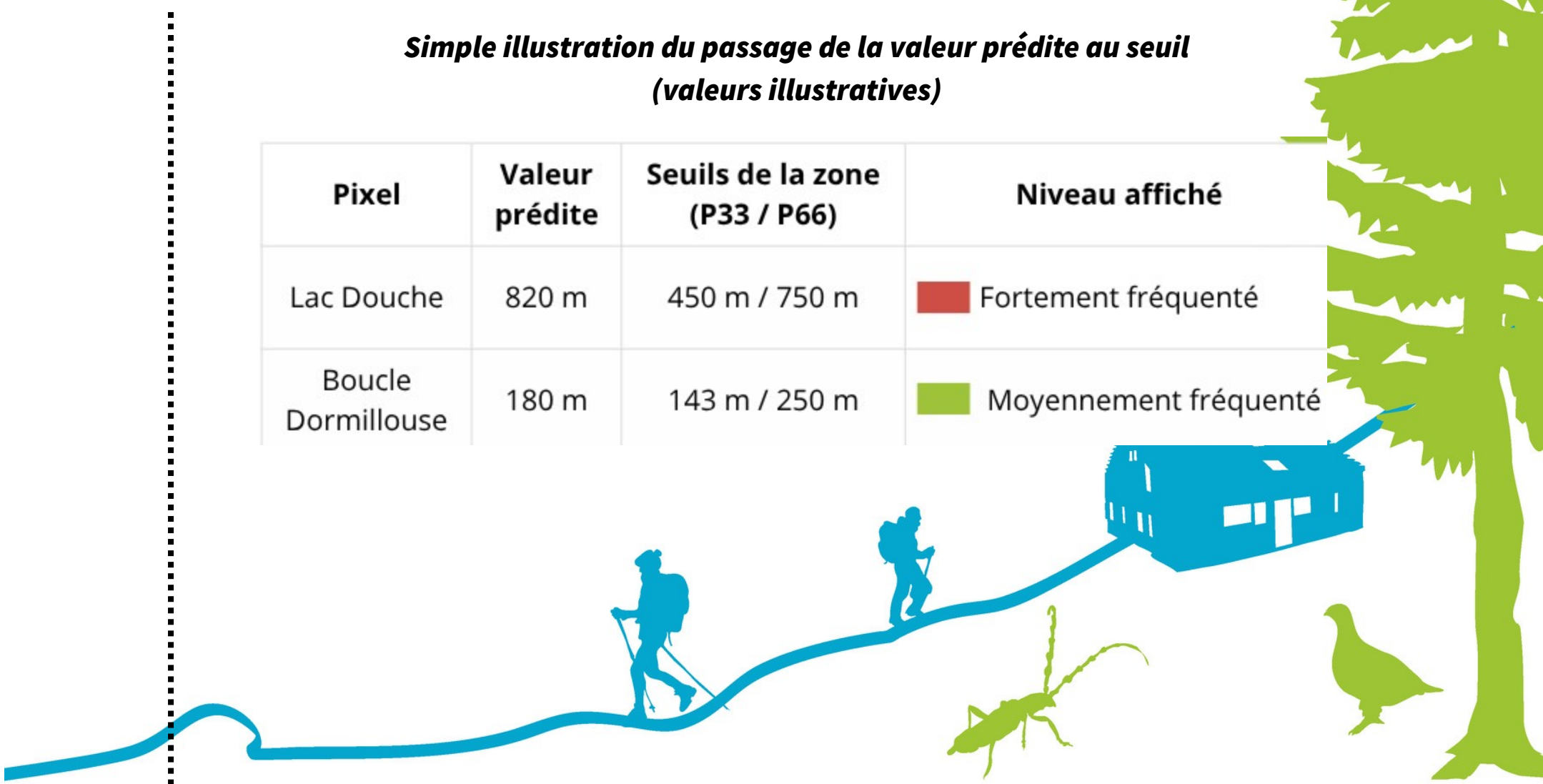


Prédictions à titre indicatif → seuils calculés selon une méthode de clustering spatial en cours de finalisation

Les seuils de fréquentation dans chaque pixel sont actuellement retravaillés, pour de meilleurs résultats

Simple illustration du passage de la valeur prédite au seuil (valeurs illustratives)

Pixel	Valeur prédite	Seuils de la zone (P33 / P66)	Niveau affiché
Lac Douche	820 m	450 m / 750 m	Fortement fréquenté
Boucle Dormillouse	180 m	143 m / 250 m	Moyennement fréquenté



De la prédiction à l'usage : l'automatisation

Dalla predizione all'uso: l'automazione

Objectif : le système (extraction des données météo et prédictions journalières à J+15) tournant seul, sans intervention humaine

Chaque jour, automatiquement :

- ➔ Récupération des données météo prévisionnelles PEARPEGE (J+4) et ECMWF (J+15)
- ➔ Lancement des prédictions pour les 15 prochains jours

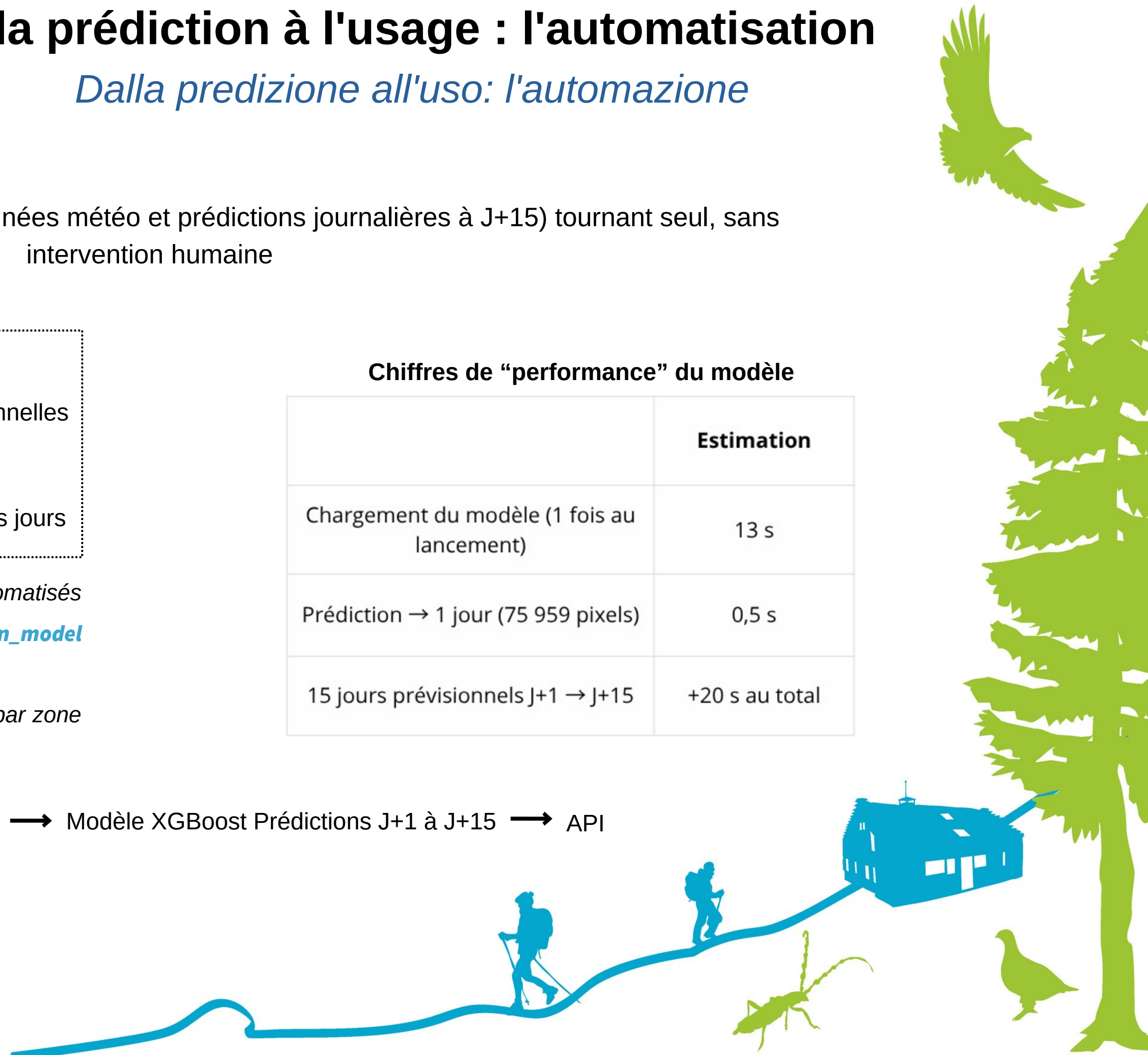
Infrastructure : *GitHub Actions - deux workflows automatisés ([fetch_weather](#) pour l'extraction des données météo & [run_model](#) pour le lancement des prédictions)*

Données statiques : *localisation des pixels, altitude, seuils par zone calculés une seule fois*

Chiffres de "performance" du modèle

	Estimation
Chargement du modèle (1 fois au lancement)	13 s
Prédiction → 1 jour (75 959 pixels)	0,5 s
15 jours prévisionnels J+1 → J+15	+20 s au total

Météo PEARPEGE/ECMWF → Modèle XGBoost Prédictions J+1 à J+15 → API



De la prédiction à l'usage : restitution

Dalla predizione all'uso: restituzione

L'API (cours de développement) → **objectif** : rendre les prédictions **accessibles** et **interrogeables** depuis une application tierce (ex : une page web)



Le widget (client) envoie une géométrie (un itinéraire) et l'API renvoie la distribution des niveaux de fréquentation captés sur les pixels intersectant cette géométrie ou le seuil dominant.

ex : 55 % fortement fréquenté, 45 % moyennement fréquenté sur l'itinéraire x

Fonctionnement de l'API (illustration) - cas du seuil dominant

Widget

renvoyé

Romarc Juvanon, CC BY-SA 4.0

GR54 Tour de l'Oisans
■ Difficile

Météo
☁ 11°C - 4 m/s - 2mm

AFFLUENCE

Demain	● Fortement fréquenté
J+2	● Moy. fréquenté
J+3	● Moy. fréquenté
.....	

API

Pixels intersectant la géométrie des seuils à J+1		distribution
● Fort	██████████	55 %
● Moyen	██████████	45 %
● Faible		0 %

géométrie + date (+ heure envisagée)

seuil dominant retourné pour le jour de prévisions "aujourd'hui"

Piste envisagée : ventiler les prédictions journalières à l'échelle horaire en s'appuyant sur des schémas de fréquentation existants



Conclusion

Conclusion

- Les conditions météo jouent un rôle déterminant dans la pratique connectée des sports de nature
Le condizioni meteorologiche svolgono un ruolo determinante nella pratica connessa degli sport all'aria aperta;
- Les traces géonumériques peuvent constituer un proxy fiable de la fréquentation réelle (cf. comparaisons traces/écocompteurs)
Le tracce geodigitali possono costituire un indicatore affidabile dell'affluenza effettiva (cfr. confronti tra tracce e contatori ecologici);
- La dimension prédictive de la fréquentation semble être utile pour anticiper et gérer les flux (un modèle entraîné sur 4 ans est capable de prédire un indicateur de fréquentation jusqu'à J+15 sur l'ensemble du réseau de sentiers du Parc national des Écrins)
La dimensione predittiva della frequentazione sembra essere utile per anticipare e gestire i flussi (un modello addestrato su 4 anni è in grado di prevedere un indicatore di frequentazione fino a J+15 sull'intera rete di sentieri del Parco Nazionale degli Écrins);
- **Mais des limites subsistent** : le biais de représentativité des traces (représentant une minorité de pratiquant·e·s, d'autant plus constituée de “super-contributeur·rice·s” : une minorité contribuant à la majorité des activités) & l'accès pérenne aux jeux de données de fréquentation pour ré-entraîner les modèles
Ma permangono alcuni limiti: il bias di rappresentatività delle tracce (che rappresentano una minoranza di praticanti, costituita per di più da “super-contributori”: una minoranza che contribuisce alla maggioranza delle attività) e l'accesso continuativo ai set di dati di frequentazione per riaddestrare i modelli.





Suites envisagées au projet par le Parc national des Écrins

Prospettive future del progetto da parte del Parco Nazionale degli Écrins:

- Expérimentation pilote de l'intégration de la prédiction dans le portail Geotrek Écrins
Sperimentazione pilota dell'integrazione delle previsioni nel portale Geotrek Écrins;
- Développement d'un widget de visualisation (horizon 2027)
Sviluppo di un widget di visualizzazione (entro il 2027)
- Publication à terme des codes sur une plateforme open source Git
Pubblicazione a termine dei codici su una piattaforma open source Git;
- Poursuite de la collaboration avec l'UGA et l'INRAE pour améliorer et actualiser le modèle
Proseguimento della collaborazione con l'UGA e l'INRAE per migliorare e aggiornare il modello

Interreg



Cofinancé par
l'Union Européenne
Cofinanziato
dall'Unione Europea

France – Italia ALCOTRA



Remerciements

Ringraziamenti

MERCI POUR VOTRE ÉCOUTE !

