

# XAI et information géographique: application aux reconstructions paléoenvironnementales

Bastien Zimmermann<sup>1</sup>, Matthieu Boussard<sup>1</sup>, Nicolas Boulbes<sup>2</sup>, Sophie Grégoire<sup>3</sup>  
<sup>1</sup> craft.ai

<sup>2</sup> Institut de Paléontologie Humaine, Fondation Albert Ier, Paris, UMR 7194 HNHP, EPCC-CERP Tautavel

<sup>3</sup> UMR 7194 HNHP, EPCC-CERP Tautavel

bastien.zimmermann@craft.ai, matthieu.boussard@craft.ai, nicolas.boulbes@cerptautavel.com,  
sophie.gregoire@cerptautavel.com

## Résumé

*Ce travail illustre les apports de l'explicabilité à la reconstruction des paléoenvironnements. Il s'agit de construire un modèle permettant, à partir des données de fouilles, de déterminer l'environnement correspondant à un niveau archéologique et une période du Paléolithique donnée. Dans ce contexte, la prédiction seule du modèle a moins de valeur que les explications sous-jacentes qui permettent aux archéologues de remettre en cause leurs hypothèses. Dû aux incertitudes sur les données, ce travail porte sur les explications orientées données, comme data-Shapley. Enfin, cet article propose l'utilisation d'un système d'information géographique permettant d'exploiter au maximum les informations issues des outils d'explicabilité.*

## Mots-clés

XAI, explicabilité, Système d'information géographique, apprentissage automatique, communautés animales, paléoenvironnements.

## Abstract

*This work shows the contribution of explicability to the reconstruction of paleoenvironments. It aims at building a model allowing, from excavation data, to infer the environment corresponding to a given layer and a given paleolithic period. In this context, the prediction of the model alone has less value than the underlying explanations that allow archaeologists to question their assumptions. Due to the uncertainties in the data, this work focuses more on data-oriented explanations tools, such as data-Shapley. Finally, a contribution of this article is the use of a Geographic Information System allowing us to exploit to the maximum the information we can obtain from the explainability tools.*

## Keywords

XAI, explainable AI, Geographic Information System, machine learning, animal communities, palaeoenvironments.

## 1 Introduction

Rendre les systèmes à base d'Intelligence Artificielle (IA) fiables est l'un des enjeux majeur de la recherche et du

développement autour des techniques d'apprentissage machine. La fiabilité d'un système est aisément remise en cause lorsque celui-ci fonctionne de façon obscure ou bien s'il repose sur de mauvaises fondations par exemple de mauvaises données. L'intelligence artificielle explicable (XAI) est l'un des piliers permettant d'aboutir à une IA de confiance. En particulier, les méthodes explicatives centrées sur les données s'inscrivent dans une démarche visant à recentrer l'attention sur les données qui ont autant, voire plus d'importance que le modèle dans le processus d'apprentissage automatique. En effet des données corrompues ou utilisées dans le mauvais contexte mènent à des conclusions erronées. Dans ce contexte, les algorithmes évaluant la qualité des points de données, permettent d'orienter plus efficacement le travail de nettoyage, d'augmentation, d'amélioration de la qualité de celles-ci et ainsi de construire des modèles plus performants. Nous avons ici utilisé **Beta-Shapley** sur des données de fouilles archéologiques, travaillées dans le cadre du programme ANR SCHOPPER [6], afin d'attribuer une valeur à nos différents points de données.

L'objectif est d'utiliser un algorithme d'apprentissage machine pour prédire un biome (une unité écologique positionnée géographiquement) en fonction des espèces animales caractéristiques de ce biome ou des écorégions qui le compose. Les restes fossiles d'animaux retrouvés dans un niveau archéologique sont utilisés pour estimer, en se basant sur le principe de l'actualisme, le biome associé et par conséquent les conditions climatiques qui avaient cours au moment de la mise en place du niveau.

Beta-Shapley permet d'avoir une analyse en amont de ce processus et d'analyser ce qui sert à entraîner notre modèle. Dans un premier temps, la pertinence de l'outil a été évaluée au travers de l'identification de données de haute et faible qualité. Nous montrons la cohérence entre la valeur du point et son impact sur la performance d'un modèle. Ensuite une analyse plus poussée met en valeur les différentes informations que peut apporter cet outil. Finalement, les explications apportées par ces outils permettent de voir nos données sous un nouveau jour et ouvrent des perspectives inédites, notamment via l'exploitation de la di-

mension spatiale. L'utilisation croissante des systèmes d'information géographique couplée à des modèles de distribution d'espèces (Ecological Niche Modelling) pour l'étude des paléoenvironnements[12], justifie le développement de méthodes explicatives d'apprentissage machine sur ce type d'outil.

Modèles prédictifs

## 1.1 Contexte

Afin d'apporter une illustration concrète des apports des outils d'XAI et plus spécifiquement les outils XAI orientés data, nous avons choisi l'exemple de la reconstruction paléoenvironnementale à partir d'un site archéologique. Ainsi, il s'agit de déterminer l'environnement et le climat correspondant à une époque donnée en considérant certains indices biologiques identifiés lors de fouilles archéologiques. Le site utilisé est la grotte paléolithique de la Caune de l'Arago à Tautavel, dans le sud de la France, qui a bénéficié de 54 ans de fouilles et d'études multidisciplinaires d'une séquence stratigraphique de 15 mètres d'épaisseur, développée entre 690 000 ans et 90 000 ans BP. [1]. Ce site a livré près de 600 000 objets répartis dans 55 niveaux archéologiques. La richesse de cet enregistrement paléolithique, la qualité de la conservation des vestiges et le système d'enregistrement standardisé des données dont il bénéficie, en font l'un des meilleurs terrains d'application des outils présentés ici pour les reconstructions paléoenvironnementales. Les reconstructions paléoenvironnementales se basent sur le principe de l'actualisme. Il se base sur l'hypothèse que les systèmes biologiques du passé fonctionnent de la même manière que ceux que l'on peut observer actuellement. Par exemple s'il est possible d'observer la répartition actuelle des différentes espèces animales pour différentes régions du globe, il est théoriquement possible de déterminer le climat du passé à partir d'un assemblage faunique fossile grâce aux affinités écologiques connues des espèces animales actuelles en les transposant aux communautés animales fossiles. Ce principe permet d'identifier le climat correspondant à une couche archéologique donnée, à partir des restes d'ossements retrouvés dans cette dernière. Toutefois, de nombreux facteurs de risques pouvant conduire à une conclusion erronée sont à prendre en considération, tels :

- la qualité du jeu de données actuelles
- la représentativité des taxons retrouvés lors des fouilles (état de conservation et biais dûs à la prédation)
- l'évolution des espèces (adaptation, migration, disparition)
- le principe même d'actualisme.

Des approches d'apprentissage machine pour la reconstruction des paléoenvironnements ont été proposées dans [10] à partir de données polliniques. Nous proposons ici d'utiliser des modèles basés sur des données fauniques. Plus fondamentalement, là où les auteurs se focalisent sur l'aspect prédictif des modèles, nous nous intéressons ici aux explications des prédictions que l'apprentissage machine peut fournir.

## 2 Présentation des Données

Les données utilisées dans cet article sont de deux natures. Les unes actuelles, utilisées comme référence pour la constitution des modèles et les autres, archéologiques avec l'objectif de les classer en utilisant des modèles d'apprentissage machine et de définir ainsi les conditions environnementales associées.

### 2.1 Les jeux de données sur les environnements actuels

Le dataset actualiste utilisé est le *wildfinder dataset* [2], c'est une représentation biogéographique de la biodiversité terrestre. L'unité de base est l'écorégion comme représenté Fig. 1. Les écorégions sont établies en fonction de critères biogéographiques, définies comme "une unité étendue de terre ou d'eau qui contient un assemblage d'espèces, de communautés naturelles et de conditions environnementales, qui se distingue au plan géographique" Chaque éco-

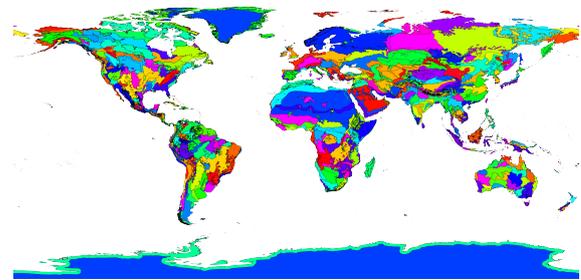


FIGURE 1 – Les différentes écorégions du monde

région contient une liste des espèces présentes sur les plus de 26000 répertoriées, ainsi qu'un type de biome associé. Ce dernier est un ensemble d'écosystèmes caractéristique d'une aire biogéographique. Chacun des 14 biomes est représentatif d'un certain climat (Fig. 2). Un niveau encore plus général de description existe, celui des écozones qui, au nombre de 8, représentent bien la répartition de la faune actuelle sur la planète. Les deux que nous considérons dans ce travail, sont le Paléarctique correspondant à l'Europe, l'Afrique du nord, les deux-tiers nord de l'Asie ainsi que le Moyen-Orient (sauf l'Arabie), et le Néarctique correspondant à l'essentiel de l'Amérique du Nord, c'est à dire les écozones de l'hémisphère nord. La richesse des données est présentée par la Fig. 3. Les différents biomes ne sont pas représentés de manière égale, il y a plus d'écorégions de *Temperate Broadleaf and Mixed Forests* que de *Montane Grasslands and Shrublands*. De plus certains biomes ont une diversité d'espèces plus faible, un nombre moins grand d'espèces est présent en moyenne dans les écorégions de la *Tundra* que ceux de *Temperate Coniferous Forests*.

Le premier outil d'explicabilité sur les données utilisable est une visualisation de celles-ci. La figure Fig.4 permet d'avoir des représentations alternatives. L'utilisation de la PCA (Principal Component Analysis) et t-SNE (t-distributed Stochastic Neighbor Embedding) permettent d'obtenir en deux dimensions une représentation pertinente

des données. En effet la PCA permet de transformer les variables entre elles et de ne garder que les deux composantes principales, décorréllées des autres et expliquant au mieux la variance. L'algorithme t-SNE, lui, a pour caractéristique de conserver la proximité des points pendant la transformation réduisant la dimension.

Ces deux représentations haut niveau permettent de constater que certain clusters existent et d'apprécier la dimension Vapnik–Chervonenkis de nos données (la théorie du même nom vise à expliquer l'apprentissage d'un point de vue statistique). En effet il semble à priori possible de distinguer les différents groupements de biomes.

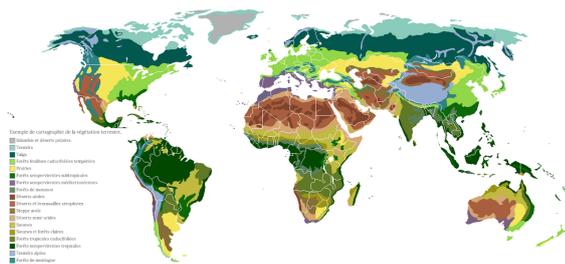


FIGURE 2 – Les différents biomes du monde

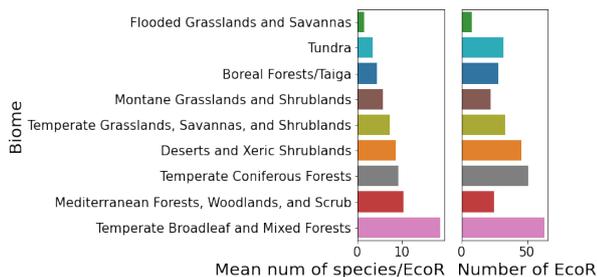


FIGURE 3 – Nombre d'espèces par écorégion par biome et nombre d'écorigions par biome. (seulement les espèces de la Caune de l'Arago)

## 2.2 Le jeu de données paléolithique

Un dataset est constitué avec les espèces de faunes identifiées dans tous les niveaux archéostratigraphiques de la Caune de l'Arago. Il rassemble les espèces de grands et petits mammifères, amphibiens, reptiles et oiseaux, déterminées à partir des restes fossiles (ossements, dents) et correspond à l'inventaire taxonomique de la communauté de vertébrés présents dans chaque couche archéologique de la Caune de l'Arago. Au total, le nombre d'espèces qui représentent les variables dans le dataset s'élève à 144. Le but de ce dataset est d'identifier les biomes et écorégions représentées dans chaque niveau archéologique afin de reconstituer l'environnement de la grotte et d'identifier le type de paysage et le climat dominant à chaque période d'occupation du site par des groupes humains.

Une fois les deux datasets constitués, nous restreignons le jeu de données WWF [2]. Géographiquement seules les écozones pertinentes pour le site de la Caune de l'Arago

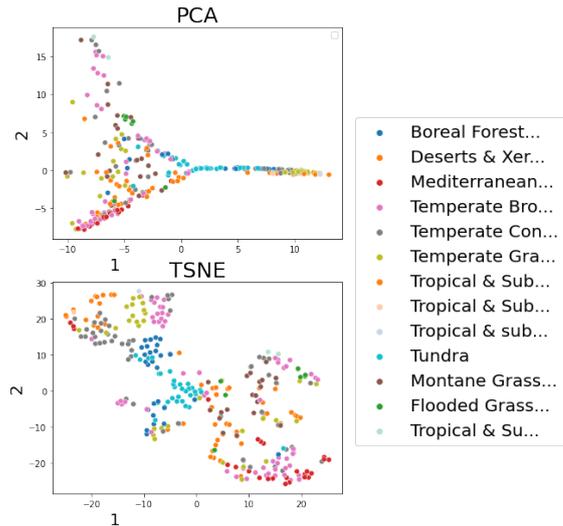


FIGURE 4 – Le jeu de donnée WWF représenté en deux dimensions

sont conservées, soit les écozones Paléarctique et Néarctique. Comme l'ensemble des taxons retrouvés lors des fouilles ne représentent qu'une petite partie des taxons existant de nos jours, le dataset WWF a dû être adapté. Les postulats de départ sont les suivants :

- Considération des espèces en présence/absence afin que l'écart entre les quantifications actuelles naturelles et celles du corpus archéologique (nécessairement plus limité) ne biaise pas les résultats des prédictions.
- Le choix de ne pas utiliser le critère d'abondance des taxons permet de plus d'éviter les biais liés à la conservation archéologique différentielle et à ceux de la sélection des espèces par leurs prédateurs (Hommes, carnivores, rapaces).
- Considération exclusive des espèces ayant été retrouvées au moins une fois dans au moins un des niveaux archéologiques, afin de ne pas prendre en compte des espèces trop éloignées des assemblages fossiles.
- Remplacement de certaines espèces éteintes par l'espèce actuelle la plus proche au plan écologique. Les espèces sans équivalent actuel (ex : le rhinocéros de prairie) sont écartés du dataset.

La figure 5 représente le dataset ainsi restreint. On peut y voir la répartition des 127 espèces restantes selon les écorégions auxquelles elles appartiennent de nos jours. Celles-ci sont plus nombreuses dans deux écorégions de l'Europe de l'Ouest (Northeastern Spain and Southern France Mediterranean forests, englobant la Caune de l'Arago, et Western European broadleaf forests) et leur densité s'affaiblit en s'en éloignant.

## 2.3 Objectif : Prédiction de types de Biome

Ces 127 taxons nous permettent donc d'entraîner un modèle dans l'objectif de prédire un biome à partir de la liste d'espèces (présence/absence) présentes dans l'écorégion. L'ob-

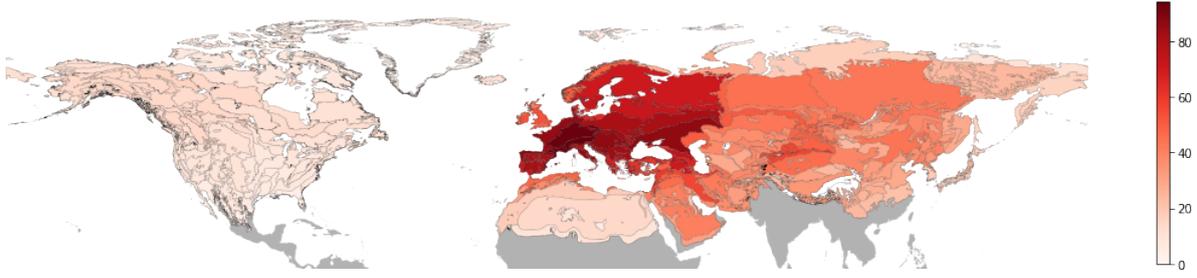


FIGURE 5 – Nombre d'espèces trouvées dans la grotte de l'Arago vivant dans chaque éco-region

jetif principal est d'exploiter ce modèle en inférence sur ces données archéologiques et prédire les conditions climatiques qui régnaient au cours de la mise en place des couches archéologiques en fonction des espèces retrouvées. Le modèle multi-classe de la figure 6 identifie les probabilités de présence des différents biomes pour chaque couche archéologique.

Cette approche prend une nouvelle dimension au travers d'outils d'explicabilité, l'inférence seule du modèle ne présentant en effet que peu de valeur pour un expert. Accompagner une prédiction d'explications permet de l'enrichir et d'apporter de nouvelles informations. Un des exemples est l'utilisation de Shap [8] pour fournir des explications accompagnant les inférences d'un modèle comme présenté figure 7.

### 3 Un outil d'explicabilité orienté données : beta-Shapley

Dans l'objectif d'attribuer une valeur aux points de donnés, A. Ghorbani et al. ont introduit le concept de data-Shapley [4]. Cette méthode se base sur le concept de théorie des jeux des valeurs de Shapley. Introduites initialement par Lloyd Shapley, elle propose une méthode équitable de répartition de gain. Ainsi, à partir d'un algorithme d'apprentissage et d'un jeu de donnée d'entraînement, data-Shapley est une métrique quantifiant la valeur de chaque point du jeu d'entraînement par rapport à la performance du prédicteur. Cette approche a de nombreux avantages, notamment que les points de faible valeur capturent les valeurs aberrantes et points corrompus, les points de haute valeur peuvent nous informer sur quel type de nouvelles données pourrait profiter à notre étude [5].

On définit la contribution marginale  $\Delta_j$  comme suit [7] :

**Definition 3.1** (Contribution Marginale). Pour une fonction  $h; j \in \llbracket 1; n \rrbracket, n = |D|$  avec  $D$  notre dataset, on définit la contribution marginale d'un point  $z^* \in D$  par rapport à  $j - 1$  points comme :

$$\Delta_j(z^*; h; D) = \frac{1}{\binom{n-1}{j-1}} \sum_{S \in D_j^{\setminus \{z^*\}}} h(S \cup z^*) - h(S)$$

avec  $D_j^{\setminus z^*} = \{S \subseteq D \setminus \{z^*\} : |S| = j - 1\}$

Le calcul de la valeur de data-Shapley pour un point de donnée est défini par [7] :

**Definition 3.2** (Data Shapley). La data-Shapley du point  $z^* \in D$

$$\psi_{shap}(z^*; U; D) := \frac{1}{n} \sum_{j=1}^n \Delta_j(z^*; U; D)$$

avec  $|D| = n; U : \cup_{j=0}^k z^j \rightarrow \mathbb{R}$  une fonction d'utilité représentant la performance d'un modèle entraîné sur un dataset  $\cup_{j=0}^k z^j; k \in \mathbb{N}$

Les data-Shapley satisfont de manière unique les propriétés suivantes [7] :

1. **Efficacité** : Les allocations s'additionnent à la valeur de l'utilité du dataset complet.

$$\forall U, \sum_{z \in D} \psi(z; U; D) = U(D) \quad (1)$$

2. **Symétrie** : Pour tout  $U$  et toute permutation  $\pi$  sur  $D$

$$\forall S \subseteq D, \psi(U(\pi(S))) = U(\pi(\psi U(S)))$$

3. **Joueur nul** : Un point  $z_i$  qui apporte une contribution marginale nulle reçoit une allocation nulle.

$$U(S \cup \{z^*\}) = U(S) \forall S \subseteq D \setminus z^* \psi(z^*; U; D) = 0$$

4. **Linéarité** :  $\forall U_1, U_2$  fonctions d'utilité,  $\forall \alpha_1, \alpha_2 \in \mathbb{R}$ ,

$$\psi(z^*; \alpha_1 U_1 + \alpha_2 U_2; D) = \alpha_1 \psi(z^*; U_1; D) + \alpha_2 \psi(z^*; U_2; D) \quad (2)$$

Une généralisation de cet outil est présentée au travers de beta-shapley définies dans [7].

Les valeurs dites beta-Shapley sont dérivées des data-Shapley en relaxant l'axiome d'efficacité (1) et ajoutant deux hyperparamètres  $(\alpha, \beta)$  décidant de poids ajoutés en fonction de la cardinalité. Une valeur élevée de  $\alpha$  mettra une importance accrue sur les ensembles de petite cardinalité et réciproquement  $\beta$  élevé mettra l'importance sur les ensembles de grande cardinalité. Une illustration des paramètres est présentée Fig. 8.

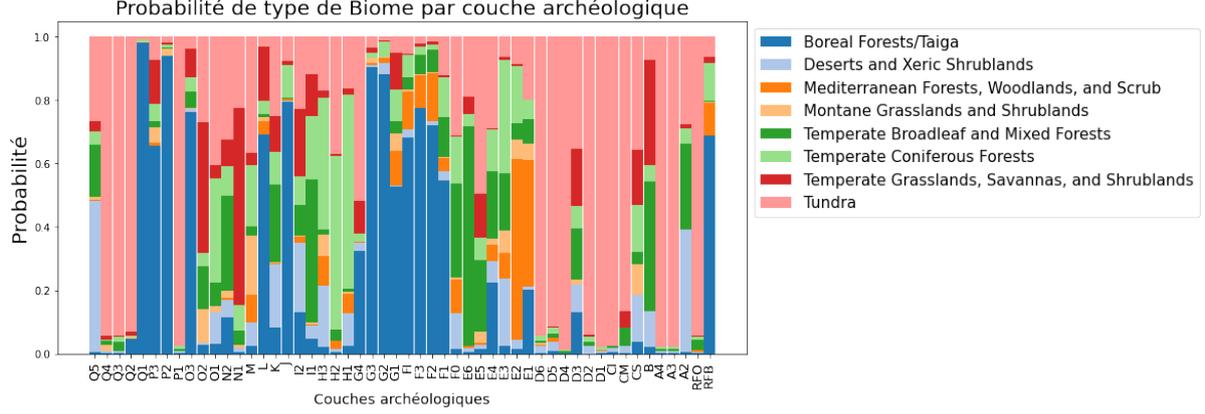


FIGURE 6 – Distribution de la représentation des biomes dans chaque couche archéologique

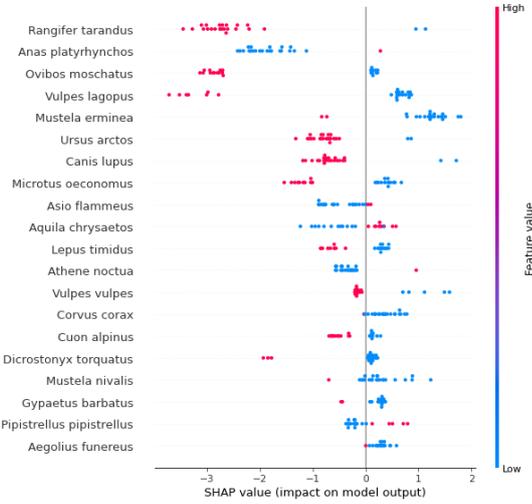


FIGURE 7 – Visualisation de l'importance des taxons via les shapley values pour le modèle de prédiction des biomes

**Definition 3.3** (Beta Shapley). La Beta Shapley du point  $z^* \in D$

$$\psi_{\text{beta}}(z^*; U; D; w^{(n)}) := \frac{1}{n} \sum_{j=1}^n w^{(n)}(j) \Delta_j(z^*; U; D) \quad (3)$$

avec  $w^{(n)} : [n] \rightarrow \mathbf{R}$  tel que :

$$n = \sum_{j=1}^n \binom{n-1}{j-1} w^{(n)}(j) \Delta_j(z^*; U; D)$$

En particulier nous utiliserons un schéma de poids défini à partir des paramètres  $\alpha, \beta$  défini comme :

$$w_{\alpha, \beta}^{(n)}(j) = n \frac{\text{Beta}(j + \beta - 1, n - j + \alpha)}{\text{Beta}(\alpha, \beta)}$$

avec :  $\text{Beta}(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta)$ ;  $\Gamma$  la fonction gamma.

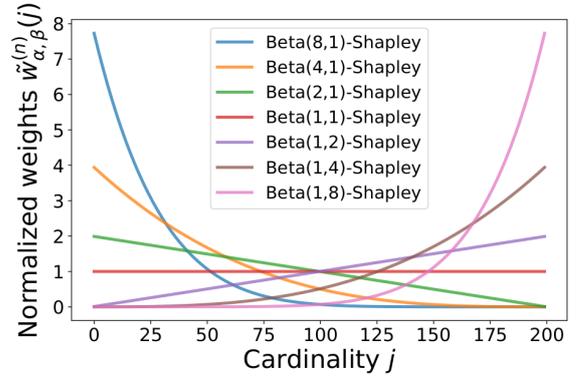


FIGURE 8 – Illustration des rôles de  $\alpha, \beta$

Pour notre étude nous avons calculé les valeurs beta-Shapley de chaque écorégion. Pour ce faire nous avons calculé la moyenne des valeurs de beta-Shapley au travers d'une stratégie de *stratified shuffle split*, nous le pouvons grâce à la propriété de linéarité (2) des valeurs beta-Shapley. Cette méthode de split permet une validation croisée en conservant le pourcentage de points de chaque classe. Cela nous permet de garder une évaluation des performances non biaisées tout en ayant une valeur pour chaque point de notre dataset.

Ainsi pour chaque split, les valeurs de beta-Shapley sont approximées par une méthode de Monte-Carlo dont la convergence est supervisée via la statistique de Gelman-Rubin [3]. Le modèle utilisé est un modèle de gradient boosting, LightGBM, et la métrique d'utilité est l'exactitude multi-classe sur l'ensemble du test correspondant à l'ensemble d'entraînement fourni par le *stratified shuffle split*.

La figure 9 donne un aperçu des résultats pour différents paramètres  $\alpha, \beta$ . Plus la valeur attribuée à une écorégion est faible plus sa couleur sera rouge, plus sa valeur sera importante plus elle sera verte, la couleur blanche correspond à 0.

On distingue aisément que  $\beta$  élevé permet d'isoler des points précis tandis qu'un  $\alpha$  élevé met en valeur des groupes de points. En effet, plus le  $\alpha$  est élevé plus on se rapproche

d'une approche du type leave-one-out qui correspond à enlever un point et mesurer la différence de performance.

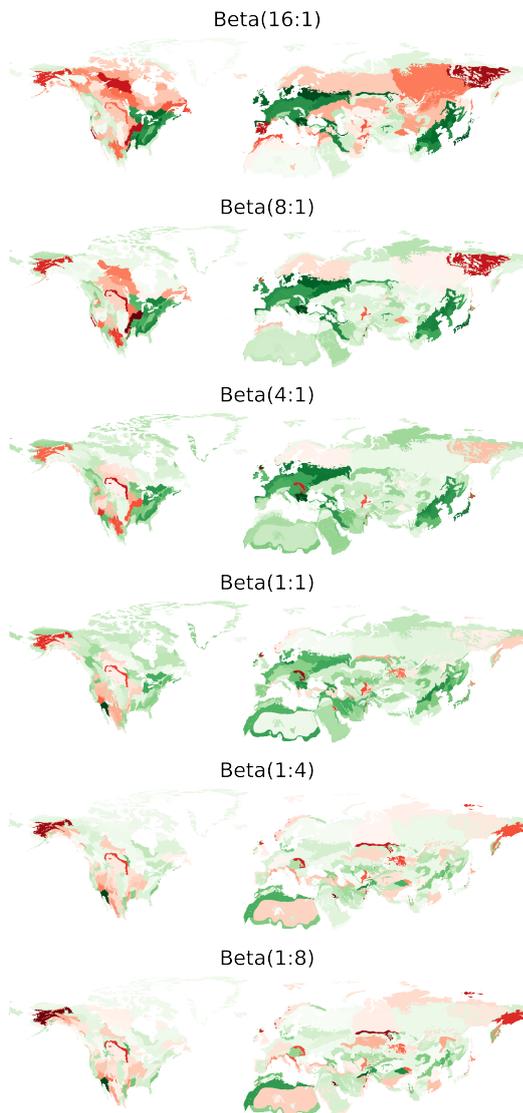


FIGURE 9 – Illustration des rôles de  $\alpha$ ,  $\beta$  sur les data-Shapley values

### 3.1 Relation entre qualité des données et Valeurs de Shapley

Notre jeu de données n'est pas parfait, en effet comme décrit précédemment nous n'avons gardé que certaines espèces, ce qui entraîne que certaines écorégions apparaissent identiques de par leur assemblage d'espèces animales; ceci est problématique car elles sont cependant caractérisées par des biomes différents. Ces écorégions qualifiées d'ambiguës sont représentées figure 10.

Les beta-Shapley values des ces écorégions sont faibles comme présenté Fig. 11. En effet celles-ci, de par leur défauts ne permettent pas un bon entraînement du modèle, voire lui nuisent et donc reçoivent une valeur faible en conséquence.

Une façon alternative de constater le fonctionnement de cet

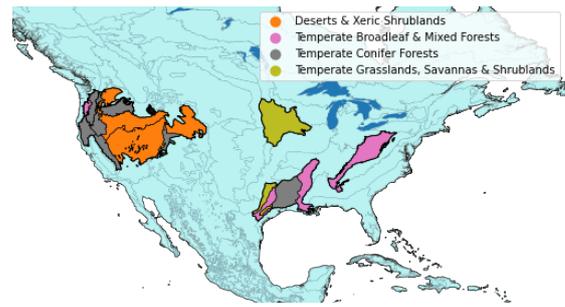


FIGURE 10 – Écorégions ambiguës

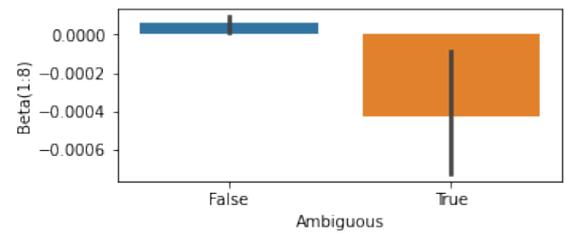


FIGURE 11 – Distribution of  $Beta(1 : 8)$  en fonction de l'ambiguïté

outil est une analyse de la correspondance performance et valeur de Beta-Shapley comme présentée Fig. 12.

Plusieurs heuristiques de valorisation de points de données sont évaluées. Chaque heuristique propose un classement des écorégions à partir duquel les points sont enlevés un à un par ordre d'importance. À chaque étape le modèle est ré-entraîné et sa performance est évaluée. L'on constate donc que beta-Shapley est significativement plus efficace qu'une méthode de sélection aléatoire. En effet, les heuristiques basées sur beta-Shapley permettent même un gain de performance en enlevant des point néfastes.

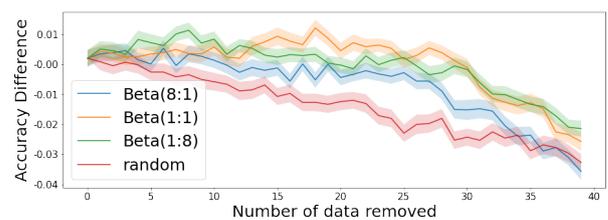


FIGURE 12 – Suppression successive de points du jeu d'entraînement et impact sur la performance

### 3.2 Inspection de tendances groupées

Les beta-Shapley paramétrées telles que  $\alpha \gg \beta$  par exemple  $Beta(8 : 1)$  mettent de larges poids sur les petites cardinalités et dé-bruïte les grandes. Par conséquent les attribution sont plus homogènes et il est possible de distinguer l'impact de groupes de points sur la performance du modèle.

Par exemple, si on groupe ces beta-Shapley values par biome, il apparait une immédiate distinction entre le biome *Temperate Broadleaf and mixed Forests* et les autres (Fig.

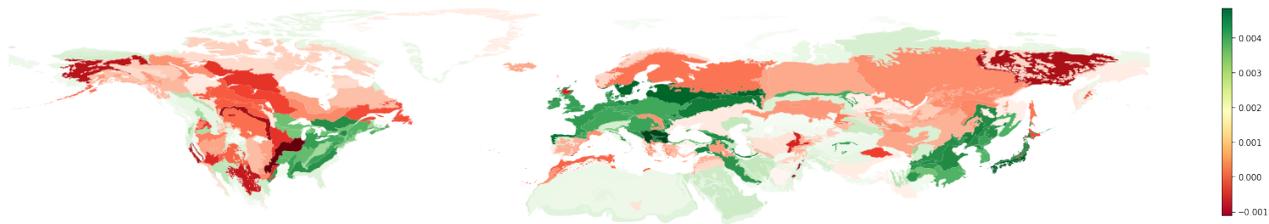


FIGURE 13 – Distribution des Beta(8 :1) par ecoregion

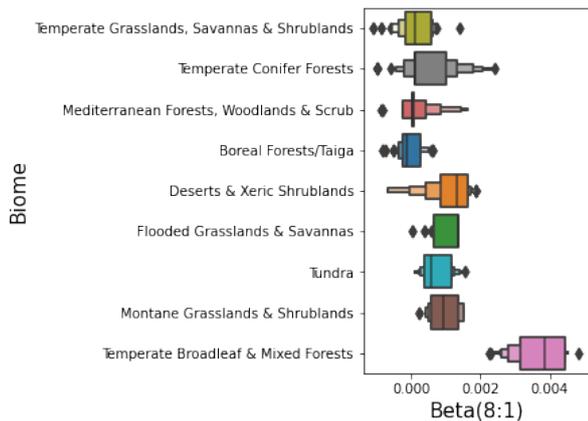


FIGURE 14 – Distribution des Beta(8 :1) par Biome

14). Les écoregions appartenant à ce Biome contribuent plus à augmenter la performance du modèle. Ce biome est à la fois le plus représenté dans le domaine de nos données (Fig 3) mais aussi, celui qui présente la plus grande variété d'espèces par écoregion. Il est possible que l'exactitude multi-classes, fonction d'utilité utilisée, ait augmenté ces valeurs en faveur du biome susmentionné. En effet celle-ci est sensible au déséquilibre des différentes classes, elle mesure mal les performances sur celles les moins représentées.

### 3.3 Inspection d'espèces influentes

Beta-Shapley nous donne accès à un classement d'écoregions reflétant leur utilité marginale en tant que point d'un ensemble d'entraînement. On considère ainsi le rang d'une écoregion le classement de sa valeur de  $Beta(1 :8)$  ordonnées dans l'ordre croissant. Ce rang permet de mettre en valeur des éléments utiles à l'entraînement du modèle. Par exemple en regardant pour les écoregions du Biome *Montane Grasslands & Shrublands* et en conditionnant sur la présence de "Buteo buteo" (Buse variable) une distinction claire apparaît (Fig 15). Pour une écoregion donnée, la présence de cet animal permet de mieux déterminer si le biome est *Montane Grasslands & Shrublands* ou non. Les écoregions contenant cet animal ont été plus informatives pour la détermination de ce biome et pour notre modèle que celles qui n'en avaient pas.

Cette information prend une nouvelle dimension une fois replacée dans son contexte. En effet une visualisation sur une carte, ou dans un système d'information géographique,

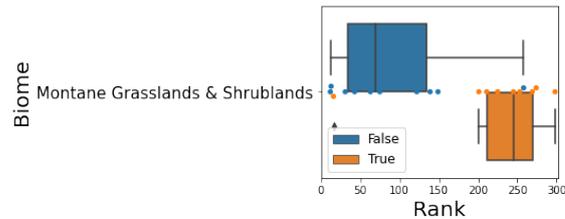


FIGURE 15 – Distribution du rang du Buteo buteo en fonction de sa présence dans le Biome *Montane Grasslands & Shrublands*

donne à un expert plus d'éléments pour arriver à des conclusions. Par exemple pour le cas présent, le Buteo buteo n'est pas présent dans les plateaux Tibétains au sud du désert du Taklamakan tandis que le biome s'étend sur une zone continue. Dans le cadre de notre étude, l'écoregion du Karakoram ouest dans laquelle vit la buse nous permet d'identifier plus facilement le biome *Montane Grasslands & Shrublands* comparé aux écoregions du même biome dans lesquelles cet animal ne vit pas.

L'on peut avoir des informations de manière similaire avec le Vulpes lagopus (renard polaire) et le biome *Temperate broadleaf & mixed forests*. Les écoregions semblant pertinentes étant : *Sarmatic mixed forests* et *Baltic mixed forests*.

## 4 Intégration à un système d'information géographique

Afin de rendre les outils d'explicabilité plus facilement utilisables, nous avons intégré les valeurs de beta-Shapley et les autres attributs disponibles dans l'outil Kepler.gl [11]. Kepler.gl est un outil de visualisation de données géospatiales open source permettant une vue interactive et en trois dimensions de nombreux calques. Ces deux éléments permettent de remettre les données dans leur contexte ainsi qu'une exploration personnalisée des données.

Les valeurs Beta(8 :1) shapley sont présentées figure 18. Chaque zone segmentée de couleur uniforme correspond à une écoregion, la couleur ainsi que la hauteur correspondent à la valeur de beta Shapley. Le minimum correspondant au bleu le plus sombre et à une hauteur nulle, le maximum au rouge le plus intense. Il est possible de constater que l'écoregion du Sahara du Nord apparaît en sur-brillance, c'est le résultat d'un clic de l'utilisateur qui peut de cette ma-

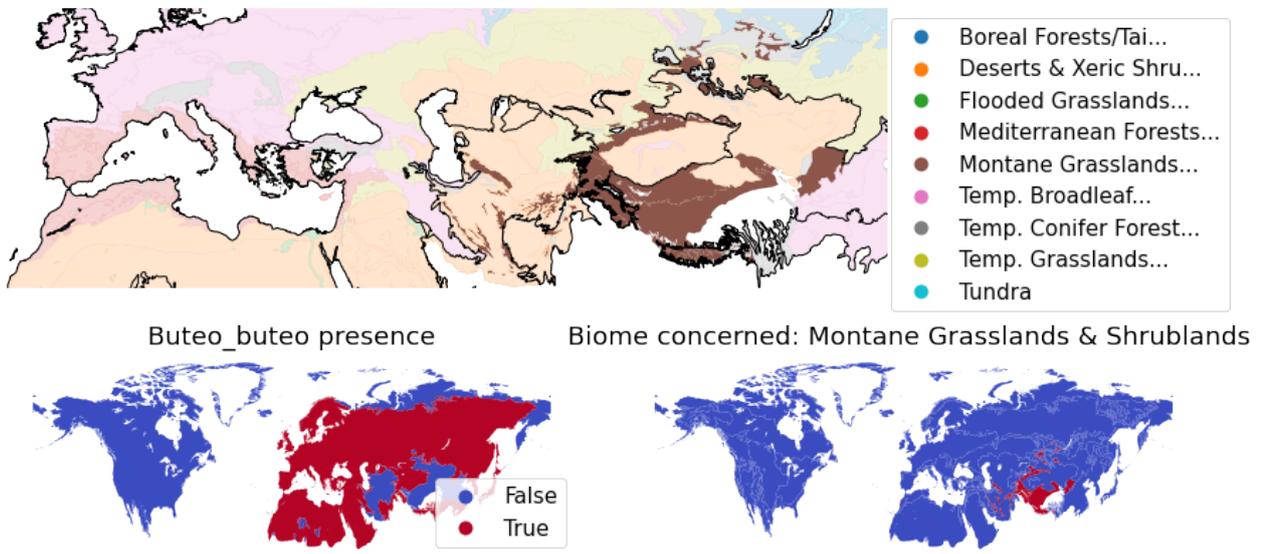


FIGURE 16 – Cartes de répartition du *Buteo buteo*

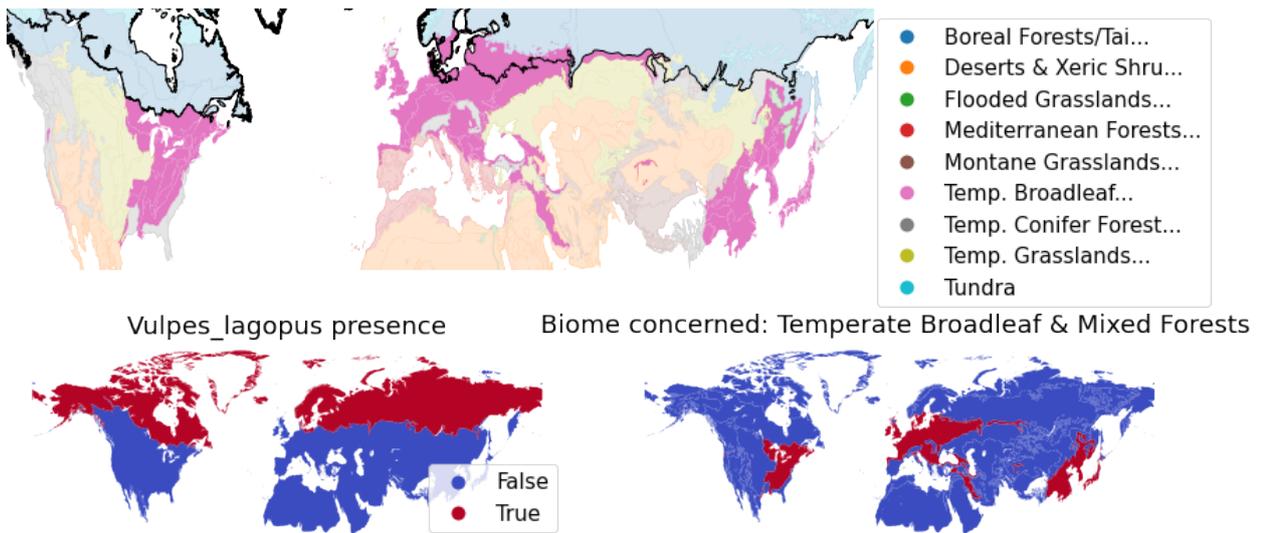


FIGURE 17 – Cartes de répartition du *Vulpes lagopus*

nière accéder au détail de ce point de données. Au delà des multiples personnalisations possibles (coloration, variable affichées, ...) l'utilisateur est libre de naviguer sur la carte et de choisir le point de vue qui lui convient le mieux dans cet espace en trois dimensions.

Les figures 19, 20 présentent les mêmes informations décrites section 3.3. Ainsi en couleur orangée apparaît le biome nous intéressant : *Montane Grasslands & Shrublands* pour la détermination duquel la présence de la buse est informative. Grâce à ces visualisations il est aisé de déterminer que l'écorégion du *North Tibetan Plateau-Kunlun Mountains alpine desert* est la seule changeant significativement de hauteur parmi les régions *Montane Grasslands & Shrublands*. Le *Buteo buteo* apparaît dans cette écorégion cependant elle possède une valeur de beta-Shapley faible ce qui contraste avec ses voisines. La forte informativité de cette écorégion peut être due à la présence ou l'absence d'un autre animal. Grâce à cette visualisation un expert du domaine disposerait à la fois d'un point de départ pour avoir un regard critique sur les données et de la nouvelle dimension que constitue les valeurs de beta-Shapley, celui de l'informativité d'un point de donnée vis-à-vis d'un modèle d'apprentissage machine.



FIGURE 18 – Exemple de l'interface de Kepler.gl

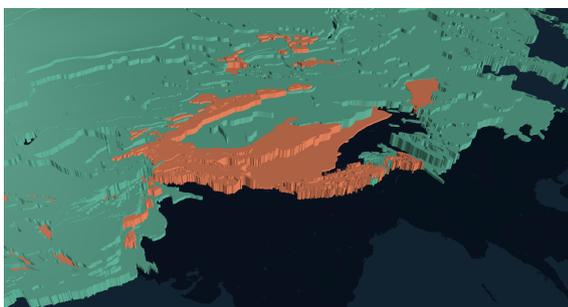


FIGURE 19 – Les informations sur *Buteo buteo* vus sur Kepler - la hauteur représente la présence de l'animal

## 5 Discussion et conclusion

Le but de cet article est de présenter l'utilisation et l'utilité d'un outil d'explicabilité centré sur les données dans



FIGURE 20 – Les informations sur *Buteo buteo* vus sur Kepler - la hauteur représente le rang

le contexte concret qu'est la reconstruction des paléoenvironnements. Nous avons montré comment l'exploitation des informations géographiques contenues dans les données permettent de mieux visualiser l'explication des modèles d'apprentissage machine. Les éléments mis en valeurs ne sont pas des conclusions, il est nécessaire qu'un paléontologue, expert du domaine, s'approprie cet outil et y apporte son savoir et ses connaissances. La fiabilité et l'utilité de cet outil qui a été décrite au fil des expérimentations reste à nuancer. En effet, l'information délivrée est complexe, l'utilité d'un point de donnée pour un modèle d'apprentissage machine par rapport à une fonction de score est pour le moins nuancée. En dépit des bonnes propriétés (section 3) beaucoup de dimensions différentes impactent les valeurs de beta-Shapley. Pour tirer des conclusions valables il est à minima nécessaire de connaître à la fois les caractéristiques du type de modèle apprenant utilisé ainsi que ses limitations mais aussi de connaître le dataset sur lequel on l'entraîne.

L'une des limitations de beta-Shapley est sa complexité. En effet le calcul des valeurs exactes est de complexité exponentielle en nombre de points. Bien que certaines méthodes permettent des approximations moins coûteuses comme les méthodes de Monte-Carlo [9], son utilisation sur des jeux de données volumineux devient quasiment impossible. Il reste cependant possible de l'appliquer à un échantillon de points représentatifs, l'utilité de cet outil serait néanmoins diminuée. Par exemple pour la détection d'anomalies, qui elle serait appropriée dans un contexte de données massives, il ne serait pas possible d'utiliser ce procédé. Il serait néanmoins possible de l'utiliser à des fins de vérification.

## Références

- [1] Henry de Lumley. *La Caune de l'Arago Tome I, Tautavel-en-Roussillon, Pyrénées-Orientales, France*. Éd. du CNRS, Paris, 2014.
- [2] World Wildlife Fund. Wildfinder : online database of species distributions, 2006.
- [3] Andrew Gelman and Donald B. Rubin. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7(4) :457 – 472, 1992.

- [4] Amirata Ghorbani and James Zou. Data Shapley : Equitable Valuation of Data for Machine Learning. *arXiv :1904.02868 [cs, stat]*, June 2019. arXiv : 1904.02868.
- [5] Amirata Ghorbani, James Zou, and Andre Esteva. Data Shapley Valuation for Efficient Batch Active Learning. *arXiv :2104.08312 [cs, stat]*, April 2021. arXiv : 2104.08312.
- [6] Sophie Grégoire, Nicolas Boulbes, Bernard Quinio, Matthieu Boussard, Caroline Chopinaud, Anne-Marie Moigne, Agnès Testu, Vincenzo Celiberti, Cédric Fontaneil, Christian Perrenoud, Anne-Sophie Lartigot Campin, Thibaud Saos, Tony Chevalier, Véronique Pois, Henry de Lumley, Marie-Antoinette de Lumley, Antoine Harfouche, Rolande Marciniack, Philippe Carrez, and Thierry Hervé. Innovative multidisciplinary method using Machine Learning to define human behaviors and environments during the Caune de l’Arago (Tautavel, France) Middle Pleistocene occupations. In Archaeopress, editor, *Big Data and Archaeology : Proceedings of the XVIII UISPP World Congress (4-9 June 2018, Paris, France), Sessions III-1 François Djindjian, (éd.) ; Paola Moscati, (éd.)*, Proceedings of the XVIII UISPP World Congress (4-9 June 2018, Paris, France), pages 28–47. 2021.
- [7] Yongchan Kwon and James Zou. Beta shapley : a unified and noise-reduced data valuation framework for machine learning. *CoRR*, abs/2110.14049, 2021.
- [8] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 4765–4774. Curran Associates, Inc., 2017.
- [9] Tomasz Pawel Michalak, Aadithya V. Karthik, Piotr L. Szczepanski, Balaraman Ravindran, and Nicholas R. Jennings. Efficient computation of the shapley value for game-theoretic network centrality. *CoRR*, abs/1402.0567, 2014.
- [10] Magdalena Sobol and S. Finkelstein. Predictive pollen-based biome modeling using machine learning. *PLOS ONE*, 13 :e0202214, 08 2018.
- [11] Open source. Kepler.gl, a powerful open source geospatial analysis tool for large-scale data sets. Online. Accessed : 2022-04-12.
- [12] Christian Willmes, Kamil Niedziółka, Benjamin Serbe, Sonja Grimm, Daniel Groß, Andrea Miebach, Michael Maerker, Felix Henselowsky, Alexander Garmisch, Masoud Rostami, Ana Mateos, Jesús Rodríguez, Heiko Limberg, Isabell Schmidt, Martin Müller, Ericson Hölzchen, Michael Holthausen, Konstantin Klein, Christian Wegener, and Georg Bareth. State of the art in paleoenvironment mapping for modeling applications in archeology-summary, conclusions, and future directions from the paleomaps workshop. *Quaternary*, 3 :13, 05 2020.