Innovative multidisciplinary method using Machine Learning to define human behaviors and environments during the Caune de l'Arago (Tautavel, France) Middle Pleistocene occupations

Sophie Grégoire[1], Nicolas Boulbes[1], Bernard Quinio[2], Matthieu Boussard[3], Caroline Chopinaud[3], Anne-Marie Moigne[4], Agnès Testu[1], Vincenzo Celiberti[1], Cédric Fontaneil[1], Christian Perrenoud[4], Anne-Sophie Lartigot Campin[1], Thibaud Saos[1], Tony Chevalier[1], Véronique Pois[1], Henry de Lumley[5], Marie-Antoinette de Lumley[5], Antoine Harfouche[2], Rolande Marciniack[2], Philippe Carrez[6], Thierry Hervé[7]

[1] HNHP- UMR 7194, Université de Perpignan, CNRS, MNHN, Centre Européen de Recherches Préhistoriques, Tautavel. Corresponding author: gregoire@cerptautavel.com

[2] CEROS – EA 4429 -Université Paris Nanterre, UFR Segmi

3 Craft Ai. WeWork, Paris

[4] HNHP- UMR 7194, Muséum national d'Histoire Naturelle, CNRS, UPVD, Centre Européen de Recherches Préhistoriques, Tautavel

[5] Institut de Paléontologie Humaine, Paris

[6] Immersion-Tools, Saint-Brieuc

[7] 16 rue Schaeffer, 93300 Aubervilliers

Abstract

SCHOPPER ANR project aim to develop a method to test research hypotheses using Machine Learning and immersives virtual environments. After 54 years of excavations, the Caune de l'Arago cave delivers a large amount of raw data and many archaeometric databases were developed through the study of this Lower Paleolithic site. The current tests focused on the evaluation of paleoenvironmental conditions and behavioral ones in each archaeological levels between 560,000 years and 90,000 years BP. All archaeological and environmental variable providing information on these issues were collected and organized by the multidisciplinary team to be studied by Machine Learning approach able to build models based on current repositories or Expert hypothesis, to learn on these models and then classify the archeological levels according to these patterns. The different steps of this innovative method are presented in this paper.

Keywords: Pleistocene, Lower Paleolithic, Machine Learning, Caune de l'Arago

Résumé

Le projet ANR SCHOPPER vise à développer une méthode pour tester des hypothèses de recherche en utilisant les techniques de Machine Learning et des environnements virtuels immersifs. Des testssur l'évaluation des conditions paléoenvironnementales et comportementales dans chaque niveau archéologique entre 560.000 ans et 90.000 ans BP. ont été menés. Toutes les variables archéologiques et environnementales fournissant des informations sur ces questions ont été collectées et organisées par l'équipe multidisciplinaire pour être étudiées selon une approche de Machine Learning capable de construire des modèles basés sur des référentiels actuels ou des hypothèses d'Experts, et de classer ensuite les niveaux archéologiques selon ces derniers. Les différentes étapes de cette méthode innovante sont présentées dans cet article.

Mots-clés : Pléistocène, Paléolithique inférieur, apprentissage automatique, Caune de l'Arago

Prehistoric archeology is a rapidly changing field thanks to the arrival of digital technologies which today allow the acquisition, management, conservation and valuation of data at undreamed performance levels ten years ago. Multidisciplinary studies, more and more frequent in the field of prehistory, are beginning to provide a multitude of data that constitute a favorable substrate for the development of research tools using Machine Learning. In recent years, works based on these tools with various applications, related to taphonomic archaeological issues (Arriaza *et al.* 2016; Egeland *et al.* 2018; Domínguez-Rodrigo and Baquedano 2018; Byeon *et al.* 2019), Geochronology (Petrelli *et al.* 2017) or environmental ones (Sobol and Finkelstein 2018; Sobol *et al.* 2019; Žliobaitė 2019), and paleolithic art (Wang *et al.* 2010) are emerging.

In terms of human behavior, researches using ML are numerous and growing. Far from our prehistoric preoccupations, they concern for example the management of situations of conflicts or disasters (Provitolo *et al.* 2015). Regarding prehistoric behavior, research using Machine Learning is still rare (Brizi Godino *et al.* 2018), because it requires the use of many examples of clearly characterized behaviors, to feed the learning phases and because archeological data are still too few, heterogeneous and hypothetical. However, the need for analysis of multidisciplinary data on Paleolithic behaviors constitute a major research issue whereas classical methodological solutions (experimental archeology) are rapidly reaching their limits.

The SCHOPPER project aims to create a digital full use of databases through an innovative system able to test research hypothesis on Paleolithic human behavior, using Machine Learning (craft ai. company), and later, through simulations in virtual environment (Immersion Tools company), with the support of researcher on Knowledge Management (CEROS team of Nanterre Univ.). This program is based on the lower Palaeolithic site of the Caune de l'Arago, in the south of France. It exploits the field data collected for more than 50 years excavations and the results of multidisciplinary studies. These archaeological and environmental data, exceptionally rich and dense, are confronted with models of environmental processes and behaviors, developed thanks to different methods and tested according to the problematic. Virtual reconstructions, partly based on the predictions obtained, then allow the researchers to immerse themselves in past environments that favors formulations and tests of hypotheses (Grégoire *et al.*, submitted). This article presents and discusses the methodology of this new numerical research tool, particularly the ML approach, and the first results obtained in reconstructing paleoenvironments and Paleolithic behaviors.

1. Material and method

1.1. Site presentation

The pilot site to develop this new research is the Caune de l'Arago cave in Tautavel, in the south of France (Figure 1), which has benefited from 54 years of excavations (Lumley *et al.* 2014). This site contains nearly 600,000 coordinated objects, from 55 archaeological layers, recorded within a stratigraphic sequence of 15 meters thick, developed between 690,000 years and 90,000 years BP. (Perrenoud *et al.* 2016; Falguères *et al.* 2015; Falguères *et al.* 2004). The richness of this Paleolithic sequence, the quality of remains conservation, the exceptional excavation duration with a standardized and continuous data recording model, make it one of the best fields of development and application of the method presented in this article.

After 54 years of data collecting and multidisciplinary studies, a large amount of raw data was acquired through excavations and pluridisciplinar studies. While a monographic work of synthesis is completed (Lumley *et al.* 2015, in progress), archaeologists have enough perspective on their data to start the next step: using Ai to explore data.

1.2. The data organization phase

Because of its multidisciplinary nature, research in Prehistory makes it difficult to cross heterogeneous, often partial and for some of them, hypothetical data. A phase of Mind Mapping

	Complexes	Ensembles stratigraphiques	❖ Grandes unités archéostratigraphiques	■ Unités archéostratigraphiques	○ Sous-unités archéostratigraphiques	* Horizons
	COMPLEXE SOMMITAL		RFB			
			RFO			
	COMPLEXE SUPÉRIEUR		A	A1		
				A2		
				A3		
				A4		
			B			
			C	Cs		
				Cm		
				Ci		
400 000 ans	COMPLEXE MOYEN	Ensemble stratigraphique III	D	Ds	Ds1	
					Ds2	
				Di	Dis	Dis3
						Dis4
					Dib	Dib5
						Dib6
			E	Es	Es1	
					Es2	
					Es3	
				Ei	Ei4	
					Ei5	
					Ei6	
			F	Fs1		
				Fm	Fm2	
					Fm3	
					Fm4	
				Fi ou Inter F/G	Fia inter F/G sup	Fia5
						Fia6
					Fib inter F/G inf	Fib7
						Fib8
			G	Gs1		
450 000 ans				Gm	Gm2	
					Gm3	
				Gi4		
		Ensemble stratigraphique II	H	H1		
				H2		
				H3		
500 000 ans			I	I1		
				I2		
			J			
550 000 ans		Ensemble stratigraphique I	K			
			L			
			M			
			N	N1		
				N2		
			O	O1		
				O2		
				O3		
			P	P1		
				P2		
				P3		
			Q	Q1		
				Q2		
				Q3		
				Q4		

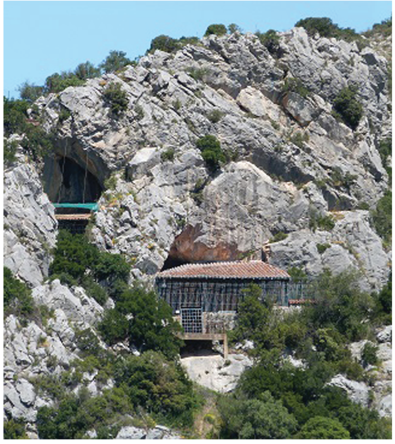

Figure 1. The Caune de l'Arago cave (Tautavel, Pyrénées-Orientales, France).

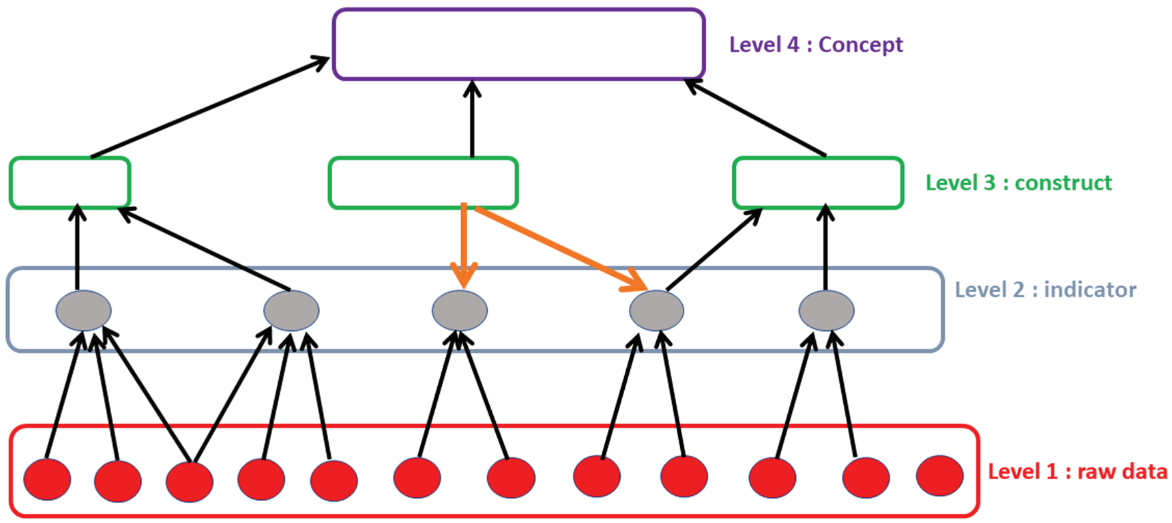

Figure 2. A conceptual model of research.

was used to structure the tacit and explicit knowledge related to each domain to hierarchize them from the initial raw data to the global concepts they allow to illustrate. Key concepts at the crossroads of disciplines were selected to develop sub-models of scenarios on which Machine Learning techniques could be applicable. The necessary and useful variables involved in a given scenario were then selected for each discipline. These data are extracted from a site-specific database called 'materiel paléontologique et préhistorique' augmented by researchers' databases. These heterogeneous data are stored in different formats and computer media.

Researchers, from management to medical study, consider there are 4 classic levels of abstraction (MacKenzie *et al.* 2005). In the case of SCHOPPER, the levels can be considered as follows (Figure 2):

- Level 1: The first-level raw data captured in the central database is the palaeontological and lithic material descriptive and archaeometric characteristics. These first-level data are precise and specific to each piece or object but they do not yet make sense.
- Level 2: These data must be aggregated and interpreted to become indicators.
- Level 3: From indicators, formative or reflective constructs (Bedford *et al.* 2017) are developed.

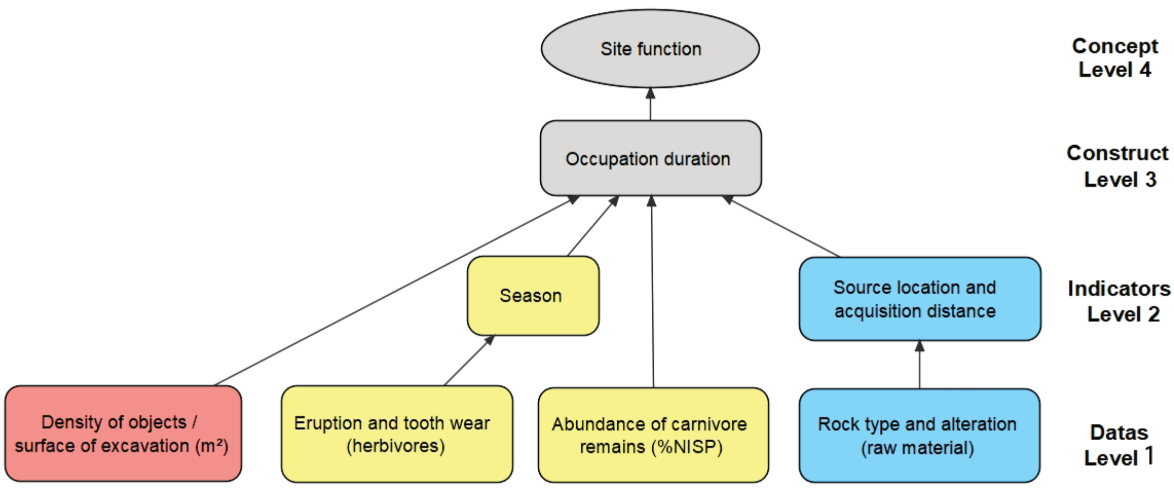

Figure 3. Example of a part of the cognitive map of the occupation duration and site function scenario (red: general archaeological considerations, yellow: faunal raw data and indicators, blue: lithics raw data and indicators, grey: constructs and concepts).

- Level 4: Finally these constructs participate to the definition of the terminal concepts which are the aim research questions which can only be treated in an interdisciplinary way.

Constructs and concepts are research hypotheses based on the analysis of multidisciplinary data and measured or calculated indicators. Cognitive maps by discipline were realized (Cosette 2008), then used to work on a construct. The elements of the map involved in the formation of a construct were put together in scenarios allowing to work on great concepts like mobility, occupation duration or site function (Figure 3).

2. Development of two methods of Machine Learning adapted to types of research hypotheses

Based on this first phase of data organizations and scenario definition, the second iterative stage led to the development of Machine Learning approaches for the generation of predictive models, ultimately allowing the testing of scientific hypotheses.

Supervised learning is a type of system in which both input and desired output data are provided. Input and output data are labelled for classification to provide a learning basis for future data processing.

In this context, two approaches appeared necessary to manage the different kinds of research hypothesis in the field of archeology.

Given that some research questions can benefit, for the learning step, of external repositories or not, we were led to develop two approaches, one for environmental questions and other for behavioral questions.

The former can rely on current ecological repositories according to the principle of actualism. The second one must necessary use prehistoric references and requires to work differently for the learning step. The principle is to use the most explicit and best characterized archeological data to classify the others thanks to algorithms tracking all the informative potential of raw data and indicators whose classical exploitation by researchers are still difficult.

These approaches could be qualified as 'exogenous method' for environmental questions using actualism principle and external repositories and as 'endogenous method' for behavioral questions using Expert Labeling. They are defined as follows:

2.1. Exogenous method: the example of biome prediction using fauna proxy

This method consists in generating predictive models learned on current and sub-current repositories (pollen spectra and open-source fauna database), following the principle of actualism, and then applied to the different levels of the Caune de l'Arago to deduce the paleoenvironments. This approach was used with two proxies related to paleoenvironments, pollen and fauna. The archeological pollen databases correspond to numerous samples throughout the entire Caune de l'Arago sequence (Renault-Miskovsky 1981; Lartigot-Campin 2007; Lartigot-Campin *et al.* in prep.). Each archaeological unit is therefore characterised by one or more pollen samples, broken down into the number of grains per taxon. Faunal proxies can also be used alone or combined with the pollen proxy. Faunal proxies is used here to describe an example of the work process.

2.1.1. Data preparation phase

Fauna database from Arago cave corresponds to a taxonomic inventory per archaeological unit of the entire community of vertebrates, large and small mammals, amphibians, reptiles and birds (Moigne *et al.* 2006; Desclaux 1992; Hanquet, Desclaux 2011; Manzano 2015; Magniez *et al.* 2013; Lebreton *et al.* 2016, 2017). The choice to work in 'presence-absence' is motivated by the quantitative heterogeneity of different groups of vertebrates and the bias due to predation on the representativity of certain species. This is particularly the case of large herbivores hunted by Paleolithics and some rodents accumulated by the action of some specialized predators.

The external repository used for the learning step, is the Wild Finder database (World Wildlife Fund., 2006). It contains presence/absence data for the world's terrestrial amphibians, reptiles, birds, and mammals, by terrestrial ecoregion which represents as many datapoints. Ecoregions are defined as 'relatively large units of land that contains a distinct assemblage of natural communities and species, with boundaries that approximate the original extent of the natural communities prior to major land use change' (Olson *et al.* 2001). In the dataset, each ecoregion is characterized by a main biome (also called major habitat type), the output class chosen.

In order to prepare a training set as compatible as possible with the Caune de l'Arago faunal data, the data from the WWF dataset have been restricted to the biogeographic realm Palearctic and to the following eight biomes: Tundra; Montane Grasslands and Shrublands; Boreal Forests/Taiga; Temperate Coniferous Forests; Temperate Broadleaf and Mixed Forests; Temperate Grasslands, Savannas and Shrublands; Mediterranean Forests, Woodlands and Scrub; Deserts and Xeric Shrublands.

A current referent is assigned to the species from Arago cave. Three cases may occur: (1) the species is still present today, (2) the species is extinct, but a current taxon with closed ecological characteristics can be substituted for it, (3) the fossil species has no equivalent today and is not taken into account. Two tests were developed: the first on the vertebrate community of the Caune de l'Arago and the second only on mammals of the site but including list of species from other Middle Pleistocene localities (presented here).

2.1.2. Exploration

Programming is carried out from the notebook interface Jupyter with the language Python. Pandas library provides data processing structure and analysis. The reference dataset has first to be explored and analysed to detect potential issues like unbalanced classes, points of data with few observations, skewed distributions, missing values etc... We then use classical multivariate analysis tools like PCA and also tools to visualize high dimensional data, such as t-SNE (van der Maaten and Hinton 2008), which improves the visualization of complex point clouds (Figure 4).

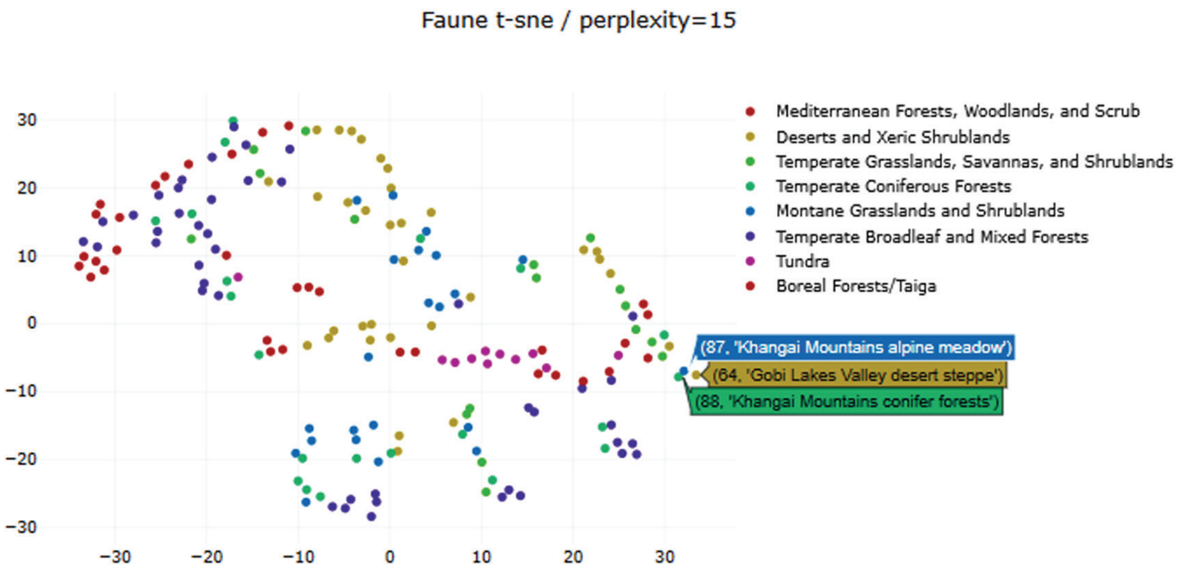

Figure 4. 2D visualization of high dimensional data with TSNE.

2.1.3. Modeling and validation

The machine learning algorithm chosen is 'Extremely randomized trees' (Geurtz *et al.* 2006) which is robust, available in the scikit learn library (Pedregosa *et al.* 2011). The modeling phase involves finding the optimal learning parameters and balancing the complexity and generalization capacity of the model. For a better suitability, two successive models are constructed: the first at the global scale of the reference frames so we can determine how the problem is difficult to classify knowing we have all taxa available. The second one, reduced to the taxa present at the Caune de l'Arago and in other localities, is used for predictions. The cross between the species from Middle Pleistocene sites including Arago cave and the WWF dataset occurs for more than 100 taxa of mammals (large mammals, lagomorphs, rodents, insectivores, bats). The validation of the models (k-fold Cross-validation with 8 fold) and the confusion matrix resulting from this step show that the classification errors are moderate (accuracy about 60%).

2.1.4. Predictions

Once the classifier is validated, it is applied to the faunal series of the different levels of the Caune de l'Arago. The predictions but also all the probabilities of belonging to each class are provided, which in this case, makes it possible to identify the biomes in competition in each level. Past communities in relation with the Pleistocene climatic variations (so-called 'non-analogous') results from biogeographical consequences, migration event and refuge (especially in mediterranean area). Biome competition predictions are regarding here as tendencies to climatic conditions and landscape, and should not be considered as aggregate or mozaic habitats.

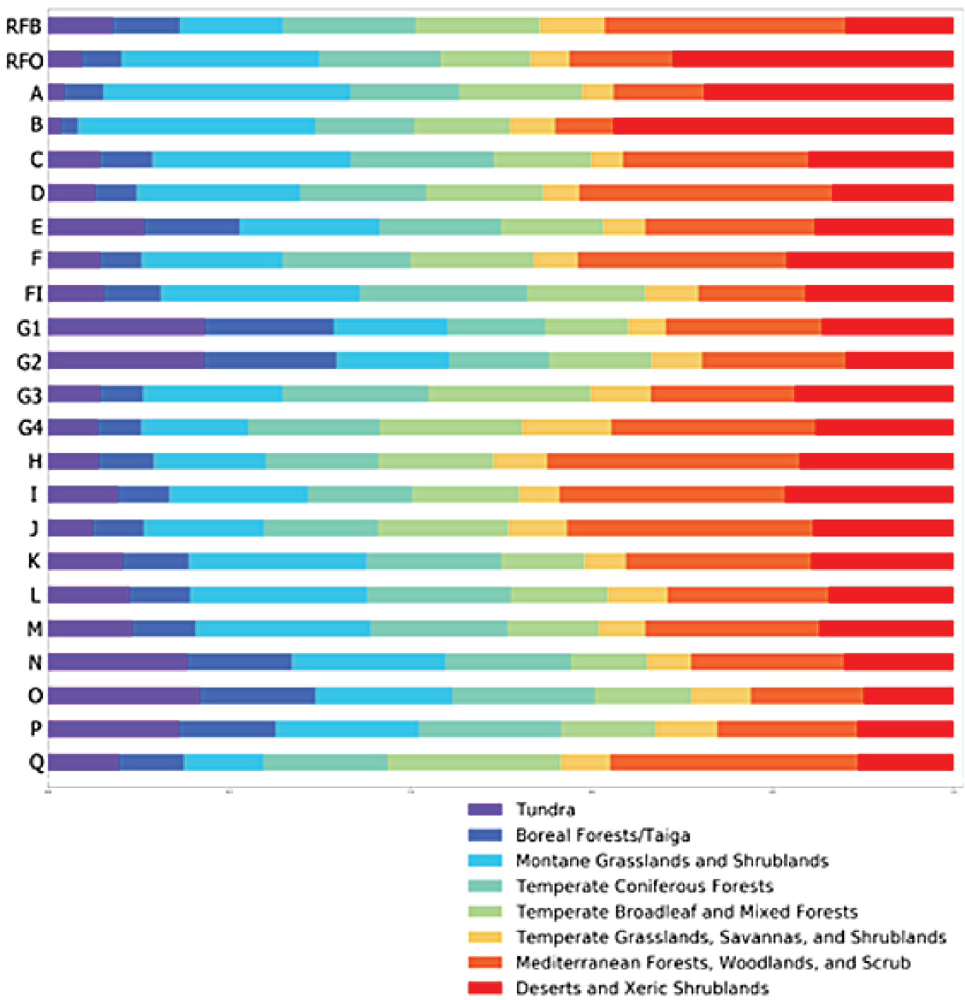

Figure 5. Biome's predictions in the stratigraphic levels of Caune de l'Arago cave.

In the detail, from the bottom to top of the stratigraphy, we can describe the following climatic cycles. It is for example interesting to show the progressive general pattern evolution of the probabilities of discriminant and opposite classes as Mediterranean and Toundra (Figure 5). The main biomes in competition are temperate in the lower level Q (Mediterranean, Temperate Broadleaf and mixed forests). The composition is radically different in the level P, O and N where the probabilities of predictions of cold environments are greater (Tundra, Montane Grasslands, Taiga). Progressively, from level M to level H the Mediterranean class becomes largely majoritary before decreasing in UA G4 and G3, but temperate biomes still remain dominant. Levels G2 and G1 mark a major change: they are characterized by high probability of the biomes Tundra (majority class) and Taiga. The other levels from the upper part of the CM III are not very contrasted and balanced between temperate and cold climatic conditions. Deserts and Xeric Shrublands are for the first time the majority class in level B, the probability of this biome is also high in next level A and RFO. Finally mediterranean habitat is the most likely environment in the RFB layer.

2.1.5. Explanation and results analysis

To interpret the predictions, explanatory algorithms such Shap (Lundberg and Lee 2017) were used. This step is crucial, it makes it possible to identify the taxa that influence the prediction of such or such biomes on a given archeostratigraphic layer, facilitates the return to the data and assists in the interpretation of the results.

First representation provides a *global interpretability* of the model: the Shap values can show how much each predictor (species) contributes, either positively or negatively, to the model. The variable importance plot lists the most significant variables: in the example of the level O, top variable as the arctic fox (*Vulpes lagopus*) has a high predictive power (Figure 6).

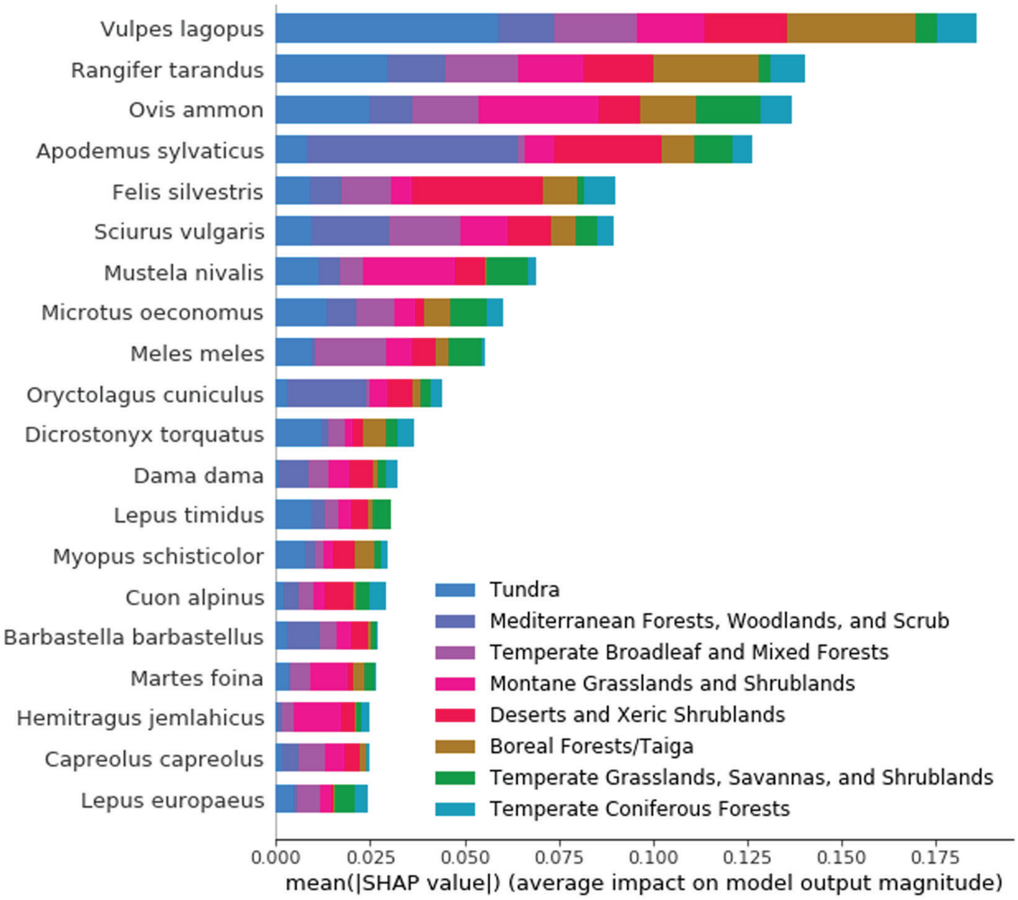

Figure 6. Variable importance plot for the level O.
The name of the species corresponds to the current taxon used as referent.

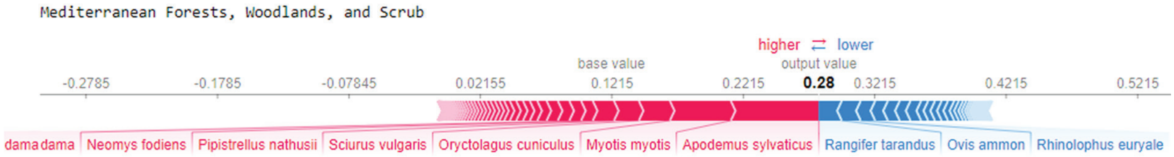

Figure 7. Visualization (force plot) of the prediction's explanation of the majority class from the level J.

The above explanation with Shap provides a *local interpretability*. It shows, for any level and any predictions, features each contributing to push the model output from the base value (the average model output over the training dataset we passed) to the model output. In the example below (Figure 7) level J belongs to the majoritary class 'Mediterranean' with a probability of 0.28 (base value p=0.1215). Features (taxa) pushing the prediction higher (in red) are the wood mouse (*Apodemus sylvaticus*), the greater mouse-eared bat (*Myotis myotis*), red squirrel (*Sciurus vulgaris*), European rabbit (*Oryctolagus cuniculus*), fallow deer (*Dama dama*). As opposite, species with a negative effect on prediction (in blue) are the reindeer (*Rangifer tarandus*), the argali (*Ovis ammon*).

Overall these results are in agreement with the climate and landscape reconstitutions based on other proxies. Especially the major cycle recognized on the Middle Complex correlated to glacial and interglacial phases (MIS 14-13-12) (Lumley *et al.* 2015; Perrenoud *et al.* 2016; Moigne *et al.* 2006). The difference of paleoenvironmental and climatic conditions within the large G unit was already demonstrated by recent study on small mammals (Lebreton *et al.* 2017). Levels G4 and G3 could correspond to a more temperate and humid climate with a more closed environment, consisting of temperate forests and Mediterranean vegetation. Levels G2 and G1 correspond to colder and open environments.

2.2. Endogenous method: the example of duration of occupation evaluation

The development of this approach is based on the observation that, to work on Paleolithic behaviors, large modernist ethnographic databases such as Lewis R. Binford one (Binford 2001) are hardly applicable as only reference in the chrono-cultural framework of the lower Paleolithic (Audouze 2013). On the other hand, it is difficult to set up benchmarks with a sufficient number of examples on fossil human population behaviors. This requires collecting examples of well-defined behaviors in publications with archaeological variables explicitly related to a particular behavior and constituting a reference dataset. The production of such a dataset, even if not easy, given the diversity of methodological approach in this field, would be necessarily subjective. So it is necessary to find a special approach to work on behavioral questions using Machine Learning.

2.2.1. Training set preparation with Expert Labeling

In this context, the exogenous method seems difficult to apply. Thereby, we have an obligation to learn from the internal database (i.e. endogenous method). Some work on the Caune de l'Arago was aimed at formulating hypotheses as to site function, occupation duration, activities, size and group composition (Lumley *et al.* 2004; Moigne *et al.* 2005; Barsky *et al.* 2005). These hypotheses were proposed on the multidisciplinary synthesis of 5 levels of occupation of the Caune de l'Arago (F, G, J, L and Inter-FG). These different occupation levels were best able to provide sufficiently contrasting archaeological data to indicate one or more behavioral changes from one level to the next. These first results constitute for the present work, a base to select the archaeological variables that affect the definition of the site function, duration and season occupation, and to qualify archaeological situations.

Once listed all the archaeological features (Table 1) that influence behaviors, the archaeologist, can propose a classification hypothesis based on an evaluation of multidisciplinary data and by using

Variables	Type	Interpretation ways	General and specific references
Density of material / surface excavation (m^2)	Numeric		Lumley *et al.* 2015
NISP lithic / (NISP lithic + fauna)	Numeric	Low density of lithic material suggests brief occupations during which predation/treatment animal ressources dominates	Rendu *et al.* 2011
Number of species of large herbivores	Numeric	Varied herbivore spectrum could refer to short-term regular hunting camps	Daujeard, Moncel 2010
Shannon's diversity index (H)	Numeric	Indice used to determine the specialization of hunting. Ability to organize massive hunts on a taxon (over a short period of time) to build food reserves (stocks) for longer consumption on a base camp. Specialised hunting could also refer to long-term residential camps	Magniez 2010; Binford 1981; David et Enloe 1993; Daujeard, Moncel 2010; Lyman 1989
Abundance of carnivores	Numeric (%)	Abundance of carnivore: repeated passage of carnivores and therefore refers to frequent movements of human groups	Aura *et al.* 2002; Daujeard, Moncel 2010; Stiner 1991; Valensi 2000; Daujeard *et al.* 2011; Costamagno *et al.* 2006
Carnivore traces	Numeric (%)		
IFO index	Numeric	Measures the intensity of bone fragmentation (NISP complete long bones / NISP long bones). Could refer specialized activities and/or partial treatment of the carcass	Brugal, Patou-Mathis 1993
Anthropic traces	Numeric (%)	Cut marks and intentional bone fracturation. Intensity of exploitation/production of animal resources	Binford 1981; Lyman 1994; Moncel, Daujeard 2010
Anatomical connections	Numeric (%)	Large number of anatomical connections in an anthropogenic level implies a partial exploitation of the carcass rather linked to brief passages of human groups. Dispersion of remains	Moigne 1983; Rendu *et al.* 2011
Season	Categorical	Rhythmicity of the growth of cervid antlers and especially the age of young herbivorous. 1 season = short or repeated at the same season; several seasons = long occupation or frequent in different seasons	Chacon *et al.* 2015
Tooth ungulate microwear	Categorial	Allows to estimate the relative duration of the mortality event(s). (i) Season or short event; (ii) event longer than a season; (iii) separated events that occurred in different non-contiguous seasons. It canbeused for expert labeling	Rivals *et al.* 2015
Manuports	Numeric (%)	A high % implies a long occupation. It could be interpreted as a raw material reserve	
Hammer	Numeric (%)	Indicate an *in situ* production rather than an introduction of finish tools. High % refers to a long occupation	
Milky quartz chaine operatoire	Categorical	Local raw material / CO complete refers to long occupation	Barsky *et al.* 2005; Grégoire *et al.* 2006; Barsky 2013
Jasper chaine operatoire		Semi-local raw material / CO incomplete refers to shorter occupation	
Miocene patinated flint chaine operatoire		Allochthonous raw material / CO incomplete refers to shorter	

Table 1. Example of variables selection to work on the question of occupation duration.

Variables	Type	Interpretation ways	General and specific references
Cortical flakes	Numeric (%)	low % indicate preparation outside and bring on site = shorter occupation high % indicates introduction and preparation on site and refers to longer occupation	Dibble 1995; Shott 1996; Clark, Barton 2017; Bicho, Cascalheira 2018
Retouching flakes	Numeric (%)	high % assumes a longer utilization of tools and rather long occupation (edges modification and sharpening and curating processes edges). *But also reused tools can suggest short occupation*	
Sharpening flakes	Numeric (%)		
Multiples tools	Numeric (%)	specialized tools could refers short occupation and multiple tools longer one	
Recycled tool or blank	Numeric (%)	high % of recycling tools, blanksor matrix, reused tool and for another function / use of available resources rather thannew supply displacement suggest long duration occupation. *But depends of the function and raw materialused*	
Ebrechures	Numeric (%)	high % of trampling so rather long occupation / fresh edges short and fast occupation	
In situ raw material	Numeric (%)	4 concentric circular areas to give the information of origin of raw material high % of *in situ* raw material (zone 1) suggest long occupation	Grégoire *et al.* 2007
Local raw material		high % of local raw material (zone 2) suggest long occupation	
Semi-local raw material		> 20% of semi-local raw material (zone 3) suggest repeated displacement related to short duration occupation	
Allochthonous raw material		> 10% of allochthonous raw material (zone 4) suggest rather short occupation	

Table 1. Continued.

his explicit and tacit knowledge. On a given case, for a specific question, the expert will propose to classify the appraised occupation layer in a category called output class or label. They are defined during a researcher multidisciplinary discussion and then tested.

2.2.2. The example of occupation duration

The concept of duration is difficult to evaluate in prehistoric sites as the mechanisms from which it arises are complex and never produce the same effects (Sullivan 1992). Yet its evaluation is crucial for the knowledge of Paleolithic behaviors (Leonova 1993; Moncel, Rivals 2011; Clark, Barton 2017; Rusch *et al.* 2019). In the case of Caune de l'Arago cave, proposals have been made for certain occupations (*op. cit.*) and serve here as a calibration for assigning an output class to the other levels of the Caune de l'Arago. Those who do not benefit from all the necessary indicators, unlike the levels that could be qualified, can thus be classified by the algorithm in one or other of the categories. The purpose of these classifications by supervised learning is here to test the levels of occupation a priori none or less explicit, based on the first assignments with the so-called classical method (i.e. without the use of ML) and to implement evidence of the variables involved in discrimination between output classes.

2.2.3. Variables choice

It depends directly on the asked question and the archaeological material analysis level. In the case of the Caune de l'Arago, these variables are multidisciplinary, quantitative or qualitative and each of them provides information in favor of one or the other class of outputs, sometimes both (or more if there are several output classes). The data can be processed in presence/absence or quantification.

A dataset is constituted with all these variables for the Caune de l'Arago 55 archaeostratigraphic layers. Output classes are ultimately assigned by experts for archaeostratigraphic layers containing sufficient arguments to propose a classification hypothesis. Learning and modeling will be done primarily on these data, then the prediction will be done on the whole dataset.

2.2.4. Exploration and modelisation

Pandas profiling (open-source package, https://pypi.org/project/pandas-profiling/) is used for exploratory and quick data analysis. This module generates a complete report for the dataset: type of variables, unique and missing values, descriptive and quantile statistics, histogram for visualizing distributions, and multiple tests of correlation.

The algorithm used is a gradient boosting classifier (Friedman, Jerome 2000). It generates a succession of small decision trees, called weak classifiers, each one just aiming at doing better than a random classifier. The combination of all those weak classifiers, produces the final decision. This algorithm is very convenient since, with few parameters, it can produce high accuracy classifier, while being robust to variable dependencies or scaling issues.

The multidisciplinary variables listed in Table 1 were selected because of their significance for the duration of occupation evaluation. Some of them are generally used for the treatment of this question, others are specific to the Caune de l'Arago according to its position within the economic territory. Depending on the situation of each of these variables in a given archeological layer, an 'Expert label' can be assigned if the data are considered sufficient. The training set thus constituted allows to generate models and a first prediction on the whole dataset.

2.2.5. Prediction

The experts labeled the dataset, where they can tell if they think the sample is short or not short. Experts can also decide not to provide any label. Furthermore even if they gave a label, they can also decide not to learn on this point, due too much uncertainty. Finally we learn only on labeled point with high expert confidence. Then we predict on all data points, labeled, labeled but uncertain, and unlabeled.

The result obtained in Figure 8 enable to compare the Expert labeling and the prediction and point the contradictory results, the similar results and the decisions taken with expert label in order to understand why. For example, the results shows: 59% of decision similar to the expert label, 9% of decision opposed to the expert label and 32% of prediction where the expert did not have enough information to decide. In these cases it is important to know which variable influence the decision or which variable is missing or insufficient to make the expert decision possible. Particularly the 9% of contradictory decisions ask questions and implies that we analyse the prediction mechanisms. An explanation phase is necessary to understand them.

2.2.6. Explanation phase and result analysis

The gradient boosting offers the possibility to show the importance of each variable in the classification (Figure 9). In this case, the first eleven features are decisives for the classification and have each and together a signification in terms of duration. The meaning of each variable is known (Table 1), but the meaning of the set of variable chosen is unknown to the expert but defined by the algorithm. The percentage of cortical flakes indicate if the reduction process takes place into the cave or not. His variation value is significant for the classification between short or not short duration in the sense that if the value is high it involves a complete and longer reduction process for flakes and tools production or on the contrary it indicates that only the using step takes place into the cave, reducing the duration of the stay. The second feature is the percentage of carnivore wears on herbivore bones, usually correlated, if it's high, with frequent movement of

Unités archéostratigraphiques	densité d'objets par surface fouillée (m2)	MP éloignée %	% Manuports	CO complète SCH9	% éclats corticaux	% éclats de retouche	% outils multiples	Indice de diversité (Shannon H)	% IFO (Indice de fragmentation osseuse)	% traces anthropiques	% traces carnivore	..	Expert label	Prediction
RFB	1613,3	11,4%	0,6%		53,1%	0,3%	1,1%	1,87	2,6%	15,2%	1,4%	...	?	NC
RFO	845	7,1%	0,3%		42,7%	0,3%	1,1%	1,88	0,0%	17,6%	0,0%	...	?	NC
A1	95	10,3%	0,0%	n	13,3%	0,0%	6,7%			0,0%	0,0%	...	?	NC
A2	182,8571429	4,4%	0,0%	n	6,5%	1,1%	3,2%	0,00	0,0%	0,0%	0,0%	...	?	C
A3	362,5	2,0%	2,0%	n	18,7%	0,4%	1,2%	0,00	0,0%	0,0%	0,0%	...	?	NC
A4	175	4,1%	2,2%	n	18,4%	3,4%	2,8%	0,00	0,0%	0,0%	0,0%	...	?	NC
B	377	3,5%	1,1%	n	33,9%	1,0%	1,4%	1,58	0,0%	1,0%	0,0%	...	?	NC
CS	445,4	2,1%	0,5%	n	48,0%	0,3%	1,3%	1,12	0,0%	3,0%	0,0%	...	NC	NC
CM	115,5555556	3,9%	0,4%	o	32,2%	0,2%	2,3%	0,60	0,0%	10,4%	0,0%	...	?	NC
NC	35,6	7,0%	4,1%	o	21,6%	0,0%	0,7%	0,69	0,0%	12,5%	0,0%	...	?	NC
D1	34,03333333	8,1%	1,7%	o	53,2%	0,4%	1,2%	1,85	0,0%	9,5%	0,0%	...	?	NC
D2	43,38095238	7,1%	0,9%	n	53,2%	0,0%	1,0%	1,75	0,0%	3,3%	2,2%	...	NC	NC
D3	30,8	10,1%	0,8%	o	51,2%	0,1%	1,2%	1,73	18,2%	10,9%	4,3%	...	NC	NC
D4	15,38461538	6,2%	1,3%	o	33,5%	0,3%	0,9%	1,59	25,0%	13,0%	6,5%	...	?	NC
D5	48,75757576	5,8%	1,1%	n	36,9%	0,2%	1,6%	1,73	7,1%	16,8%	3,2%	...	C	NC
D6	40,02777778	6,2%	0,9%	n	34,7%	0,3%	1,0%	1,60	0,0%	10,1%	1,1%	...	C	NC
E1	47,21428571	4,0%	3,6%	o	24,8%	0,5%	1,6%	1,36	0,0%	15,2%	1,3%	...	NC	NC
E2	145,0714286	4,8%	3,1%	o	24,1%	1,0%	2,3%	1,46	1,2%	12,1%	0,7%	...	C	C
E3	63,76744186	6,0%	5,1%	o	23,9%	1,9%	3,5%	1,55	2,9%	14,8%	0,8%	...	NC	NC
E4	27,63636364	7,1%	4,3%	o	25,7%	2,7%	3,2%	1,29	0,0%	10,9%	1,5%	...	NC	NC
E5	15,13333333	4,5%	3,1%	n	27,4%	2,0%	2,3%	1,30	0,0%	19,6%	2,2%	...	C	C
E6	28,10869565	4,0%	8,1%	n	24,7%	1,7%	1,7%	1,37	0,0%	16,2%	2,4%	...	NC	NC
F0	16,18181818	2,7%	3,7%	n	23,8%	0,0%	2,7%	0,99	0,0%	26,3%	1,4%	...	C	C
F1	299,1276596	4,8%	4,1%	o	20,5%	3,1%	3,2%	1,28	0,5%	21,7%	2,3%	...	NC	NC
F2	569,9791667	5,3%	5,7%	o	21,1%	2,6%	2,6%	1,24	0,9%	21,5%	2,2%	...	NC	NC
F3	450,1020408	5,1%	7,6%	o	23,6%	2,6%	2,2%	1,29	1,4%	20,3%	3,0%	...	NC	NC
FI	74,79	2,9%	4,2%	n	25,5%	1,7%	2,4%	1,49	0,4%	18,5%	2,9%	...	NC	NC
G1	706,5961538	2,4%	7,0%	o	25,5%	2,1%	1,9%	1,75	1,8%	21,7%	2,9%	...	NC	NC
G2	866,1071429	3,8%	10,0%	o	21,9%	2,6%	2,7%	1,73	1,7%	21,6%	2,8%	...	NC	NC
G3	699,5333333	4,8%	9,8%	o	19,8%	3,2%	3,4%	1,49	0,9%	24,5%	3,3%	...	NC	NC
G4	580,3934426	7,2%	8,1%	o	22,9%	5,0%	3,4%	1,42	0,8%	23,9%	3,1%	...	NC	NC
H1	52,48076923	10,4%	7,1%	n	23,4%	4,4%	5,6%	1,48	0,0%	25,0%	2,5%	...	NC	NC
H2	41,73469388	10,2%	8,9%	o	22,6%	2,3%	3,1%	1,67	3,1%	31,0%	3,0%	...	NC	NC
H3	35,02	8,3%	4,6%	n	23,4%	1,4%	6,4%	1,78	0,0%	30,7%	4,1%	...	NC	NC
I1	69,13461538	5,5%	5,3%	n	13,1%	1,3%	2,3%	1,64	2,0%	32,0%	3,8%	...	C	NC
I2	24,94339623	7,3%	3,5%	n	23,6%	1,0%	3,4%	1,32	0,0%	31,2%	3,5%	...	NC	NC
J	285,056338	6,0%	1,3%	n	17,4%	0,7%	1,8%	1,19	0,3%	30,3%	3,5%	...	C	C
K	26,25	3,5%	1,8%	n	10,9%	1,1%	2,1%	1,56	1,6%	28,1%	8,9%	...	C	C
L	129,9142857	3,5%	3,5%	n	15,8%	0,7%	0,6%	0,65	10,9%	23,1%	13,0%	...	C	NC
M	21,5	2,1%	5,8%	n	18,8%	0,5%	2,4%	1,47	2,9%	31,3%	7,7%	...	C	C
N1	14,47297297	5,3%	5,5%	n	42,9%	1,9%	0,0%	1,18	7,7%	16,1%	9,5%	...	NC	NC
N2	24,94666667	6,2%	1,3%	n	50,0%	1,6%	0,8%	1,45	17,4%	12,1%	7,5%	...	NC	NC
O1	9,25974026	7,3%	1,8%	n	68,8%	3,8%	1,3%	1,67	0,0%	16,4%	4,5%	...	NC	NC
O2	8,064935065	13,3%	0,9%	n	76,4%	0,0%	3,6%	1,33	0,0%	17,1%	1,2%	...	C	C
O3	6,987012987	6,8%	0,9%	n	63,1%	2,7%	1,8%	1,41	0,0%	15,9%	9,8%	...	NC	NC
P2	32,72151899	2,2%	0,9%	n	54,4%	0,8%	1,8%	1,66	3,4%	17,1%	7,4%	...	C	NC
P3	108,5375	3,5%	0,6%	o	52,0%	0,9%	1,6%	1,36	1,7%	20,6%	6,4%	...	NC	NC
Q1	178,6811594	4,7%	0,6%	o	39,0%	1,6%	1,3%	1,48	1,9%	22,8%	4,7%	...	NC	NC
Q2	40,53571429	4,0%	0,3%	n	35,1%	1,6%	2,5%	1,24	7,7%	19,6%	4,1%	...	NC	NC
Q3	90,2962963	5,6%	0,2%	o	38,4%	2,8%	2,6%	1,11	0,0%	22,3%	3,1%	...	NC	NC
Q4	35,73333333	5,1%		n	22,8%	13,7%	2,2%	0,69	0,0%	9,5%	9,5%	...	?	NC

Figure 8. Result of prediction classifying by short (C, *green*), not short (NC) output class. The training set has learnt on three expert labels: short, not short, and unknown (?) assigned and predicted on the whole layers (*red: concordant predictions*). The figure presents only a selection of variables. The original dataset contains 27 variables.

Feature ranking

1. feature % éclats corticaux (0.362637)
2. feature % traces carnivore (0.232044)
3. feature % outils multiples (0.219555)
4. feature CO complète SCH9 (0.121895)
5. feature MP locale % (0.053803)
6. feature MP in situ % (0.006419)
7. feature Richesse spécifique (0.003072)
8. feature % éclats de ravivage (0.000481)
9. feature densité dobjets par surface fouillée (m2) (0.000078)
10. feature % éclats de retouche (0.000013)
11. feature % outils réutilisés (0.000002)
12. feature % connexion anatomique (0.000000)
13. feature % ébréchures (0.000000)
14. feature % Manuports (0.000000)
15. feature NR lithique / (NR lithique+faune) (0.000000)
16. feature % traces anthropiques (0.000000)
17. feature % IFO (Indice de fragmentation osseuse) (0.000000)
18. feature MP semi-locale % (0.000000)
19. feature MP éloignée % (0.000000)
20. feature CO complète QL9 (0.000000)
21. feature % Percuteurs (0.000000)
22. feature Indice de diversité (Shannon H) (0.000000)
23. feature NR Carnivores/NR Herbivores (0.000000)
24. feature CO complète SJ (0.000000)
25. feature % Pierres retouchées/Total Pierres (0.000000)
26. feature NR Herbivores (0.000000)

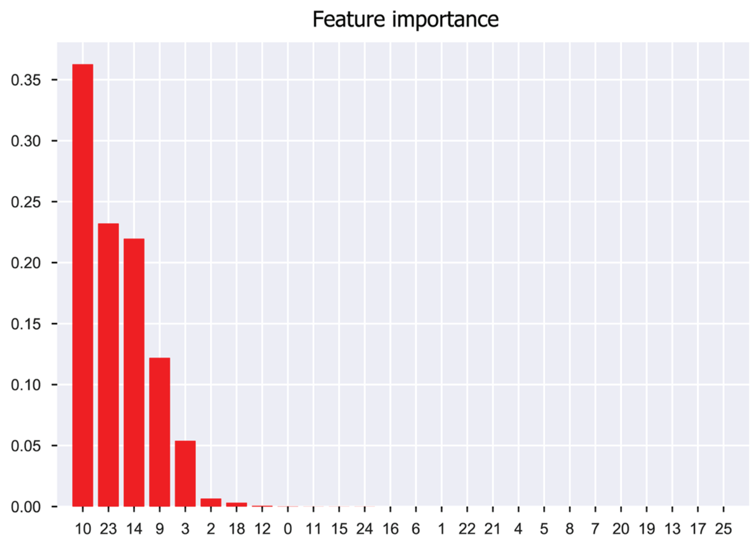

Figure 9. Feature ranking and feature importance for the prediction of short and not short occupation duration.

human groups and then with rather short-term occupation. The high value of the multiple tools seems to indicate a long duration since one blank could be optimized thanks to two or more active parts and having a long life and successives uses, and so on....

D5, D6, I1, L and P2 layers are predicted in opposition with the expert label. In each of them the expert had found indicators of short-duration, particularly on lithics indicators, considered less significant by the algorithm prediction. L is typically an equivocal layer because it is little extended, clearly stratigraphically delimited by steril deposits, with indicators of short duration: carnivores action, low percentages of cortical flakes, retouching and sharpening flakes, seasonality hunting events during early winter on reindeer (Magniez *et al.* 2011), and microwear pattern (Rivals *et al.* 2015). Occupation was interpreted as short hunting camp (Lumley *et al.* 2004). Similar patterns (especially specialized hunting) are sometimes considered as a long term occupations (Deaujard, Moncel 2010; Betts, Friesen 2004). The prediction point out this ambiguity and leads to find new indicators to be more precise or to revise the thresholds determining the limit from one class to another.

2.3. For Exogenous and Endogenous method, a new way of learning was observed

Since the implementation of the regular tests, a triple learning loop adapted from that proposed by (Argyrys and Shon 1996) have been observed. As a reminder, the Agyrys and Schon organizational learning model describes two learning loops that make it possible to learn from an error or an offset between the expected result and the result obtained. The first is a simple adaptation to the difficulty encountered, it is fast and relies on routines. The second loop requires in-depth action that challenges the usual action patterns.

In the case of the Schopper project, the simple loop is similar to that described above (Figure 10). The data-scientist and the researcher will correct outliers of the data set or modify a parameter and redo the algorithm. The second loop concerns actions that challenge the data used by looking for other internal or external data. This loop can also lead to changing the nature of the algorithm used. Finally, the third loop questions the proposed conceptual infrastructure by questioning knowledge maps or scenarios. Indeed, if the results obtained by prediction appears aberrant or confused, it can lead to reconsider the output classes and to adjust them according to the degree of resolution reached on a given problematic. For example, the output classes to characterize the site occupation duration was initially: short, long or frequent and repeated. After a first series of

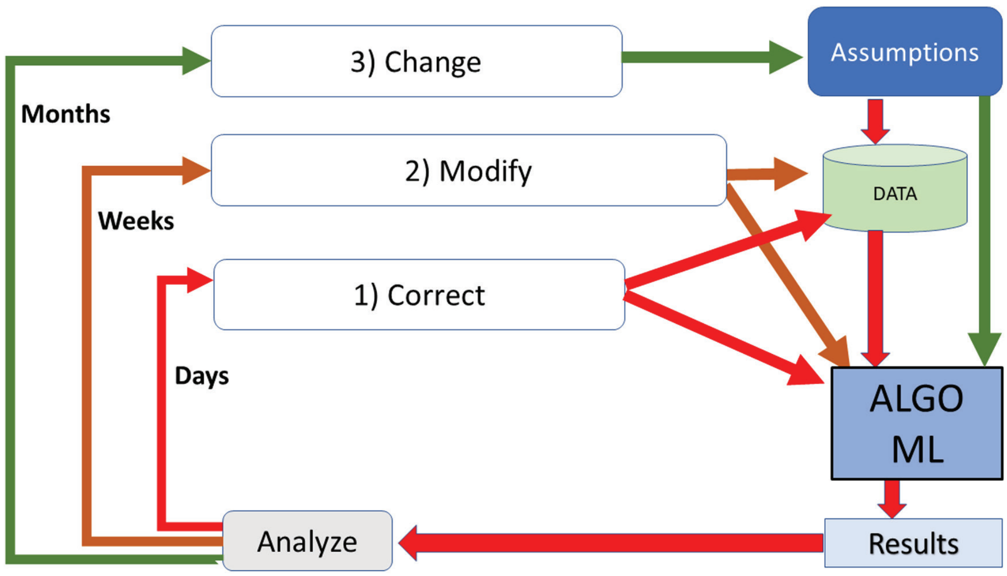

Figure 10. Principle of the triple loop learning in AI.

tests, the algorithm highlighted the unsuitability of the 'frequent' output class which corresponds to repeated occupations that can be confused with short. These first tests led to the conclusion that on the question of the occupation duration, the potentially identifiable results are: 'short' and/or 'long', 'short' or 'not short', or 'long' or 'not long' and the third output class was dropped to optimize learning and predictions. This approach led researchers to question the causal links that are usually used in their scientific reasoning and to identify new ones.

3. Discussion and conclusion

These first predictions demonstrated the potential of the Machine Learning algorithms both on environmental issues (exogenous method) and on behavioral ones (endogenous method). Their contribution is different depending on the type of hypothesis tested and the available repositories. This methodology requires a rigorous step of data preparation and the scientific care that must be provided on it. It requires the use of a new working method that combines the conceptual approach (conceptual model) with the analysis of data (Machine Learning).

The exogenous method allows to obtain usable results able to be published if they are interpreted and put in perspective as a part of the treated problematic. The example of biome reconstructions using actuals faunal repositories shows that it is possible to predict for the Caune de l'Arago archeostratigraphical layers the environment around the cave and consequently enable to reconstruct past landscapes and environmental and climatic conditions. These kinds of results can be reached thanks to other more specialized palaecological approaches for example dietary and ecological niche reconstitution (Rivals, Lister 2016) or studies based on separate different communities (Stoetzel, Montuire 2016; Magniez, Boulbes 2014; Hanquet *et al.* 2018 and references therein). The ML approach offers the possibility to generate automatically and quickly with the support of a robust frame of reference taking into account thousands of global data from different vertebrates communities. In this case the results are obtained without having recours of the expert labeling and strictly from a calculation and can be considered as more objectives. Moreover, once the training data set is done, it's possible to test many archeological faunal spectra, from different sites and periods and obtain automatised predictions with systematic explanations. In order to enrich the results, the prediction can be done with combining faunal proxies and others environmental ones, like pollen or geological variables. This method can deal with a broad spectrum of environmental issues with training and modeling on Biomes, Ecoregions, Koppen climate classes... The repositories could also be improved by others Pleistocene contemporaneous available datas.

The endogenous method, using 'Expert labeling' allows the algorithm to compare the values of the Expert labels together and to explain the indicators that contributed the most to the algorithm classification and more implicitly to the Expert labeling. It is then possible to return to these indicators to study the particular role they played in the decision-making. In this method, the explanation phase is the most important benefit of the prediction rather than the result of the prediction itself. The identification of the variable weights help the archeologists to better organise his scientific reasoning and hypothesis and give strength to the future predictions. The prediction brings raw results unusable directly but interesting to highlight the significance and the importance of the data in resolving a given question. The successive iterations, enriched at each stage by an improvement of the dataset, taking into account the previous results, progressively allows to obtain a robust prediction. Above all, this iterative work process provide at each step the necessary explanation to demonstrate the path followed to reach the result. At this stage, this method provides the most important short-term development perspectives. For the occupation duration, new tests are in progress by Expert labeling with the result of tooth ungulate microwears (Rivals *et al.* 2015) and new datasets are prepared on a given great concept like 'site function', combining archeological repository (compilation of bibliographic data) and Expert labeling on a hybrid 'exogenous-endogenous' method.

References

Argyrys, C. and Schon, D.A. 1996. *Organizational Learning: A Theory of Action Perspective*, Reading, Mass.: Addison Wesley, second edition, 344 p.

Arriaza, M.C. and Domínguez-Rodrigo, M. 2016. When felids and hominins ruled at Olduvai Gorge. A machine learning analysis of the skeletal profiles of the non-anthropogenic Bed I sites. *Quaternary Science Reviews* 139: 43-52.

Audouze, F. 2013. Les bases de données de Lewis R. Binford accessibles sur le serveur de la Maison de l'archéologie et de l'ethnologie René-Ginouvès. *Bulletin de la Société préhistorique française* 110(2): 353-355.

Aura, J.E., Bonilla, V.V., Ripoll, M.P., Valle, R.M. and Calatayud, P.G. 2002. Big Game and Small Prey: Paleolithic and Epipaleolithic Economy from Valencia (Spain). *Journal of Archaeoogical Method and Theory* 9: 215-268.

Barsky, D. 2013. The Caune de l'Arago stone industries in their stratigraphical context. *Comptes Rendus Palevol* 12(5): 305-325.

Barsky, D., Grégoire, S. and Moigne, A.M. 2005. *Variabilité des types d'occupation et d'exploitation de territoires méditerranéens entre 600.000 ans et 300.000 ans*, in M.-H. Moncel and J.L. Monnier (eds) *Données récentes sur les premiers peuplements en Europe. Actes du colloque international, (Rennes 2003)*: 565-576. British Archeological Reports International Series 1364, Oxford: Archaeopress.

Betts, M.W. and Friesen, T.M. 2004. Quantifying hunter-gatherer intensification: a zooarchaeological case study from Arctic Canada. *Journal of Anthropological Archaeology* 23: 357-384.

Bicho, N. and Cascalheira, J. 2018. The use of lithic assemblages for the definition of short-term occupations in hunter-gatherer prehistory, in A. Picin and J. Cascalheira (eds) *Short-term Occupations in Paleolithic Archaeology. Interdisciplinary Contributions to Archaeology*. Springer.

Binford, L.R. 2001. *Constructing Frames of Reference: an Analytical Method for Archaeological Theory Building using Hunter-Gatherer and Environmental Data Sets*. Berkeley: University of California Press, 563 p.

Binford, L.R. 1981. *Bones: Ancient men and modern myths*. I ed., New York: Academic Press.

Bedford, D.S. and Spekle, R.F. 2017. *Construct Validity in Survey-Based Management Accounting and Control Research, Management Accounting Section (MAS) Meeting*. <https://ssrn.com/abstract=3011357>.

Briz i Godino, I., Ahedo, V., Alvarez, M., Pal, N., Turnes, L., Santos, J.I., Zurro, D., Caro, J. and Galan, J.M. 2018. Hunter-gatherer mobility and technological landscapes in southernmost South America: a statistical learning approach. *Royal Society Open Science* 5: 180906.

Brugal, J.P. and Patou-Mathis, M. 1993. L'assemblage osseux de l'abri des Canalettes, in L. Meignen (ed.) *L'abri des Canalettes. Un habitat moustérien sur les grands Causses (Nant, Aveyron. Fouilles 1980-1986)*: 77-87. Paris : Éd. du CNRS 10.

Byeon, W., Domínguez-Rodrigo, M., Arampatzis, G., Baquedano, E., Yravedra, J., Maté-González, M.A. and Koumoutsakos, P. 2019. Automated identification and deep classification of cut marks on bones and its paleoanthropological implications. *Journal of Computational Science* 32: 36-43.

Chacon, M.G., Bargallo, A., Gabucio, M.J., Rivals, F. and Vaquero, M. 2015. *Chapter 12. Neanderthal behaviors from a spatio-temporal perspective: an interdisciplinary approach to interpret archaeological assemblages*, in N.J. Conard and A. Delagnes (eds) *Settlement Dynamics of the Middle Paleolithic and Middle Stone Age*vol. IV: 253-294. Tübingen: Tübingen Publications in Prehistory.

Clark, G.A. and Barton, C.M. 2017. Lithics, landscapes & la Longue-durée – Curation & expediency as expressions of for ager mobility. *Quaternary International* 450: 137-149.

Cossette, P. 2008. La cartographie cognitive vue d'une perspective subjectiviste: mise à l'épreuve d'une nouvelle approche, *M@n@gement* 11(3): 259-281.

Costamagno, S., Meignen, L., Beauval, C., Vandermeersch, B. and Maureille, B. 2006. Les Pradelles (Marillac-le-Franc, France): A Mousterianreindeerhunting camp? *Journal ofAnthropological Archaeology* 25: 466-484.

Daujeard, C., Moncel, M.H., Rivals, F., Fernandez, P., Aureli, D., Auguste, P., Bocherens, H., Crégut-Bonnoure, E., Debard, E. and Liouville, M. 2011. *Quel type d'occupation dans l'ensemble F de Payre (Ardèche, France)? Halte de chasse spécialisée ou campement de courte durée? Un exemple d'approche* *multidisciplinaire*, in F. Bon, S. Costamagno and N. Valdeyron (eds) *Haltes de chasse en Préhistoire. Quelles réalités archéologiques? Actes du colloque international (Touluse, 13-15 mai 2009). P@lethnologie* 3: 77-101.

Daujeard, C., Moncel, M.H. 2010. On Neanderthal subsistence strategies and land use: A regional focus on the Rhone Valley area in southeastern France. *Journal of Anthropological Archaeology* 29: 368-391.

David, F. and Enloe, J.G. 1993. L'exploitation des animaux sauvages de la fin du Paléolithique moyen au Magdalénien, in *Exploitation des animaux sauvages à travers le temps, XIIIes Rencontres d'Archéologie et d'Histoire d'Antibes, IVe colloque international de l'Homme et l'animal, Société de recherche interdisciplinaire (Antibes-Juan-les-Pins, 15-17 oct. 1992)*: 29-46. Antibes-Juan-les-Pins: Éd. APDCA.

Desclaux, E. 1992. Les petits vertébrés de la Caune de l'Arago à Tautavel (Pyrénées-Orientales). Biostratigraphie, paléoécologie et taphonomie. *Bulletin du Musée d'Anthropologie Préhistorique de Monaco* 35: 35-64.

Dibble, H. 1995. Middle paleolithic scraper reduction: background, clarification and review of the literature to date. *Journal of Archaeological Method and Theory* 2: 299-36.

Domínguez-Rodrigo, M. and Baquedano, E. 2018. Distinguishing butchery cut marks from crocodile bite marks through machine learning methods. *Scientific Reports* 8(1): 5786.

Egeland, C.P., Domínguez-Rodrigo, M., Rayne Pickering, T., Menter, C.G. and Heaton, J.L. 2018. Hominin skeletal part abundances and claims of deliberate disposal of corpses in the Middle Pleistocene, *PNAS* 115(18): 4601-4606.

Falguères, C., Shao, Q., Han, F., Bahain, J.-J., Richard, M., Perrenoud, C., Moigne, A.-M. and Lumley, de H. 2015. New ESR and U-series dating at Caune de l'Arago, France: A key-site for European Middle Pleistocene. *Quaternary Geochronology* 30: 547-553.

Falguères, C., Yokohama, Y., Shen, G., Bischoff, J.L., Ku, T.L. and Lumley, H. 2004. New U-series dates at the Caune de l'Arago, France. *Journal of Archaeological Science* 31: 941-952.

Friedman, J. 2000. Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics* 29.

Geurts, P., Ernst, D. and Wehenkel, L. 2006. Extremely randomized trees. *Machine learning* 63(1): 3-42.

Grégoire, S., Moigne, A.M., Barsky, D. and Lumley, de H. 2007. Gestion et sélection des ressources au sein d'un territoire. Un exemple de comportement économique au Paléolithique inférieur dans le sud de la France: 1725-1727, BAR International Series 1364.

Grégoire, S., Barsky, D. and Byrne, L. 2006. The Caune de l'Arago: an example of middle Pleistocene flint exploitation (Tautavel, Pyrénées-Orientales, France), in *Stone Age, Mining Age. Der Anschnitt, Beiheft 19, 2005-VIII International Flint Symposium (Bochum, 13-17 September 1999)*: 99-113. Bochum: Deutsches Bergbau-Museum.

Hanquet, C., Stoetzel, E., Desclaux, E. and Bailon, S. 2018. Etat de la recherche sur les micromammifères et l'herpétofaune quaternaires en France, in *La Préhistoire de la France*: 99-102. Paris: Hermann.

Hanquet, C. and Desclaux, E. 2011. Analyse paléoécologique des communautés de micromammifères de la Caune de l'Arago (Tautavel, France) dans le contexte des migrations de faunes en Europe méridionale au cours du Pléistocène moyen. *Quaternaire* 22(1): 35-45.

Lartigot, A.S. 2007. *Taphonomie pollinique en grotte de sédiments détritiques et de spéléothèmes: Potentiels et limites pour la reconstitution de l'environnement végétal de l'homme préhistorique sur le pourtour nord-ouest méditerranéen. Application aux sites de la Caune de l'Arago (Tautavel, Pyrénées-Orientales), de la Baume Bonne (Quinson, Alpes-de-haute-Provence), de la grotte du Lazaret (Nice, Alpes-Maritimes) et de la grotte italienne de la Basura (Toirano, Ligurie)*. Thèse de doctorat, MNHN, Paris, 545 p.

Lartigot-Campin, A.-S. *et al.* 2019. Prediction of past biomes from palynological data in archaeological context using Machine Learning: the sequence of Caune de l'Arago (Middle Pleistocene, South of France). In preparation.

Lebreton, L., Desclaux, E., Hanquet, C., Cuenca-Besco, G., Moigne, A.M., Perrenoud, C. and Grégoire, S. 2017. Variations paléoenvironnementales au sein de l'Unité Archéostratigraphique G (UA G) de la Caune de l'Arago (Tautavel, France). Apport des paléocommunautés de rongeurs. *Quaternaire* 28(3): 313-321.

Lebreton, L., Desclaux, E., Hanquet, C., Moigne, A.M. and Perrenoud, C. 2016. Environmental context of the Caune de l'Arago Acheulean occupations (Tautavel, France). New insights from micro vertebrates in Q-R levels. *Quaternary International* 411: 182-192.

Leonova, N.B. 1993. *Criteria for estimating the duration of occupation at Paleolithic sites. An example from Kamennaia Balka II*', in O. Soffer and N.D. Praslov (eds) *From Kostienki to Clovis. Upper Palaeolithic-Paleo-Indians Adaptations*: 149-157. New York: Plenum Press.

Lumley, de H., Fontaneil, C., Grégoire, S., Batalla, G., Caumon, G., Celiberti, V., Chevalier, T., Deguillaume, S., Fournier, A., Lumley, de M.A., Magniez, P., Moigne, A.M., Notter, O., Perrenoud, C., Pois, V., Pollet, G. and Testu, A. 2015. *Caune de l'Arago Tome VI. Tautavel-en-Roussillon, Pyrénées-Orientales, France: Individualisation des unités archéostratigraphiques*. Paris: Éd. du CNRS, 641 p.

Lumley, de H. (ed.) 2014. *La Caune de l'Arago Tome I, Tautavel-en-Roussillon, Pyrénées-Orientales, France*. Paris: Éd. du CNRS, 432 p.

Lumley, de H., Grégoire, S., Barsky, D., Batalla, G., Bailon, S., Belda, V., Briki, D., Byrne, L., Desclaux, K., El Guenouni, K., Fournier, A., Kacimi, S., Lacombat, F., Lumley, de M.-A., Moigne, A.M., Moutoussamy, J., Paunescu, C., Perrenoud, C., Pois, V., Quilès, J., Rivals, F., Roger, T. and Testu, A. 2004. Habitat et mode de vie des chasseurs paléolithiques de la Caune de l'Arago (600,000-400,000 ans). *L'Anthropologie* 108: 159-184.

Lundberg, S.M. and Lee, S.I. 2017. A unified approach to interpreting model predictions, in *31st Conference on Neural Information Processing Systems (NIPS 2017)*. Long Beach, CA, USA.

Lyman, R.L. 1994. *Vertebrate Taphonomy*. Cambridge: Cambridge University Press.

Lyman, R.L. 1989. Taphonomy of Cervids Killed by the 18 May 1980 Volcanic Eruption of Mount St. Helens, Washington, in R. Bonnichsen and M.H. Sorg (eds) *Bone Modification*: 149-167. Orono: Center for the Study of the First Americans, University of Maine.

MacKenzie, S.B., Podsakoff, P.M. and Jarvis, C.B. 2005. The Problem of Measurement Model Mis specification in Behavioral and Organizational Research and Some Recommended Solutions. *Journal of Applied Psychology* 90(4): 710-730.

Magniez, P. 2010. *Etude paléontologique des Artiodactyles de la grotte Tournal. Etude taphonomique, archéozoologique et paléoécologique des grands Mammifères dans leur cadre biostratigraphique et paléoenvironnemental*. Thèse de Doctorat, Perpignan: Université de Perpignan, 916 p.

Magniez, P., Moigne, A.M. and Lumley, de H. 2011. First reindeer exploitation by Acheulean people in South of France: Case study of the Lower Palaeolithic site of the Caune de l'Arago (Tautavel, Pyrénées-Orientales, France), in *Congrès International 'Deer and People' (Lincoln, UK, 10 September 2011)*. Abstract volume: 40.

Magniez, P., Moigne, A.M., Testu, A. and Lumley, de H. 2013. Biochronologie des Mammifères quaternaires. Apport des Cervidae du site Pléistocène moyen de la Caune de l'Arago (Tautavel, Pyrénées-orientales, France). *Quaternaire* 24(4): 477-502.

Magniez, P. and Boulbes, N. 2014. Environment during the Middle to Late Palaeolithic transition in southern France: The archaeological sequence of Tournal cave (Bize-Minervois, France). *Quaternary International* 337: 43-63.

Manzano, A. 2015. *Les amphibiens et les reptiles des sites du Pléistocène moyen et supérieur du pourtour méditérranéen (Caune de l'Arago, grotte du Lazaret, Baume Moula-Guercy). Etude d'herpétofaunes et reconstitutions paléoclimatiques et paléoenvironnementales*. Thèse de Doctorat, Perpignan: Université de Perpignan, 560 p.

Moigne, A.M. 1983. *Taphonomie des faunes quaternaires de la Caune de l'Arago, Tautavel*. Thèse 3ème cycle, Paris VI/Museum National d'Histoire Naturelle, 342 p.

Moigne, A.M., Palombo, M.R., Belda, V., Heriech-Briki, D., Kacimi, S., Lacombat, F., Lumley, de M.A., Moutoussamy, J., Rivals, F., Quiles, J. and Testu, A. 2006. Les faunes de grands mammifères de la Caune de l'Arago (Tautavel) dans le cadre biochronologique des faunes du Pléistocène moyen italien. *L'Anthropologie* 110(5): 788-831.

Moigne, A.M., Grégoire, S. and Lumley, de H. 2005. Les territoires de chasse et d'exploitation des matières premières des hommes préhistoriques de la Caune de l'Arago entre 600,000 ans et 400,000 ans, in J. Jaubert and M. Barbaza (eds) *Territoires, déplacements, mobilité, échanges durant la préhistoire. Actes des 126e Congrès national des sociétés historiques et scientifiques (Toulouse, 2001)*. Paris: Ed. du CTHS.

Moncel, M.H. and Rivals, F. 2011. On the question of short-term Neanderthal site occupations: Payre, France (MIS 8-7), and Taubach/Weimar, Germany (MIS 5). *Journal of Anthropological Research* 67(1): 47-75.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P. and Kassem, K.R. 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience* 51(11): 935-938.

Pedregosa, F. *et al.* 2011. Scikit-learn: Machine learning in Python. *Journal of machine learning research* 12: 2825-2830.

Perrenoud, C., Falguères, C., Moigne, A.M., Testu, A., Magniez, P., Boulbes, N., Lebreton, L., Hanquet, C., Desclaux, E., Lartigot-Campin, A.S., Celiberti, V., Grégoire, S., Lumley, de H., Viallet, C., Chevalier, T., Lumley, de M.A., Vialet, A., Fontaneil, C., Pois, V., Hu, H.M., Shen, C.C. and Michel, V. 2016. Diversité des occupations humaines en contexte glaciaire à la Caune de l'Arago, in *Colloque AFEQ CNF-INQUA Q10 (Bordeaux, 16-18 février 2016)*. Bordeaux.

Petrelli, M., Bizzarri, R., Morgavi, D., Baldanza, A. and Perugini, D. 2017. Combining machine learning techniques, microanalyses and large geochemical datasets for tephrochronological studies in complex volcanic areas: New age constraints for the Pleistocene magmatism of central Italy. *Quaternary Geochronology* 40: 33-44.

Provitolo, D., Dubos-Paillard, E., Verdière, N., Lanza, V., Charrier, R., Bertelle, C. and Aziz-Alaoui, M.A. 2015. Les comportements humains en situation de catastrophe: de l'observation à la modélisation conceptuelle et mathématique. *Cybergeo: European Journal of Geography*, Systèmes, Modélisation, Géostatistiques, document 735. <http://journals.openedition.org/cybergeo/27150> [Accessed October 26, 2019].

Renault-Miskovsky, J. 1981. Etude palynologique du remplissage de la Caune de l'Arago à Tautavel. Signification chronologique, paléoclimatique et palethnographique des flores, in H. de Lumley and J. Labeyrie (eds) *Datations absolues et analyses isotopiques en Préhistoire-Méthodes et limites*: 253-258. Paris: CNRS, Prétirage.

Rendu, W., Bourguignon, L., Costamagno, S., Meignen, L., Soulier, M.-C., Armand, D., Beauval, C., David, F., Griggo, C., Jaubert, J., Maureille, B. and Park, S.J. 2011. Approche interdisciplinaire et réflexions méthodologiques sur la question des haltes de chasse au Moustérien, in Fr. Bon, S. Costamagno and N. Valdeyron (eds) *Haltes de chasse en Préhistoire. Quelles réalités archéologiques?, Actes du colloque international (Touluse 13-15 mai 2009). P@lethnologie* 3: 61-76.

Rivals, F. and Lister, A.M. 2016. Dietary flexibility and niche partitioning of large herbivores through the Pleistocene of Britain. *Quaternary Science Reviews* 146: 116-133.

Rivals, F., Prignano, L., Semprebon, G.M., Lozano, S. 2015. A tool for determining duration of mortality events in archaeological assemblages using extant ungulate microwear. *Scientific Reports* 5: 17330.

Rusch, L., Grégoire, S., Pois, V. and Moigne, A.M. (accepted). Neanderthal and Carnivore Occupations in Unit II from the Upper Pleistocene Site of Ramandils Cave, (Port-la-Nouvelle, Aude, France). *Journal of Archaeological Science: Reports* 102038 PII S2352-409X(19)30374-8.

Shott, M.J. 1996. An exegesis of the curation concept. *Journal of Anthropological Research* 52: 259-280.

Sobol, M.K. and Finkelstein, S.A. 2018. Predictive pollen-based biome modeling using machine learning. *PLoS ONE* 13(8): e0202214.

Sobol, M.K., Scott, L. and Finkelstein, S.A. 2019. Reconstructing past biomes states using machine learning and modern pollen assemblages: A case study from Southern Africa. *Quaternary Science Reviews* 212: 1-17.

Stiner, M.C. 1991. Food procurement and transport by human and non-human predators. *Journal of Archaeological Science* 18: 455-482.

Stoetzel, E. and Montuire, S. 2016. Les rongeurs, indicateurs des paléoclimats: application aux assemblages de trois sites du nord de la France. *Quaternaire* 27(3): 227-238.

Sullivan, A. 1992. Investigating the Archaeological Consequences of Short-Duration Occupations. *American Antiquity* 57(1): 99-115.

Toniolo, A.M. 2009. Le comportement: entre perception et action, un concept à réhabiliter. *L'Année psychologique* 2009/1, 109: 155-193.

Valensi, P. 2000. The archaeozoology of Lazaret Cave (Nice, France). *Interntional Jorrnal of Osteoarchaeology* 10: 357-367.

Van der Maaten, L.J.P. and Hinton, G.E. 2008. Visualizing High-Dimensional Data Using t-SNE, *Journal of Machine Learning Research* 9: 2579-2605.

Wang, J.Z., Ge, W., Snow, D.R., Mitra, P. and Giles, C.L. 2010. Determining the Sexual Identities of Prehistoric Cave Artists Using Digitized Handprints: A Machine Learning Approach, in *Proceedings of the ACM International Conference on Multimedia (Florence, Italy, ACM, October 2010)*: 1325-1332.

World Wildlife Fund 2006. WildFinder: Online database of species distributions. <www.worldwildlife.org/WildFinder>[ver. Jan-06].

Zliobaite, I. 2019. Concept drift over geological times: predictive modeling baselines for analyzing the mammalian fossil record. *Data mining and knowledge discovery* 33(3): 773-803.