Modéliser et comprendre les campagnes de la Somme romaine à l'aide des systèmes d'information géographique

Modelling and Understanding the Countryside of the Somme in the Roman Era with Geographic Information Systems

Nicolas Revert

Édition électronique

URL : <https://journals.openedition.org/revuehn/1274>

DOI : 10.4000/revuehn.1274

ISSN : 2736-2337

Éditeur

Humanistica

Référence électronique

Nicolas Revert, « Modéliser et comprendre les campagnes de la Somme romaine à l'aide des systèmes d'information géographique », *Humanités numériques* [En ligne], 3 | 2021, mis en ligne le 01 mai 2021, consulté le 15 juillet 2021. URL : <http://journals.openedition.org/revuehn/1274> ; DOI : <https://doi.org/10.4000/revuehn.1274>

Les contenus de la revue *Humanités numériques* sont mis à disposition selon les termes de la Licence Creative Commons Attribution 4.0 International.

Modéliser et comprendre les campagnes de la Somme romaine à l'aide des systèmes d'information géographique

Modelling and Understanding the Countryside of the Somme in the Roman Era with Geographic Information Systems

Nicolas Revert

Résumés

Une des fonctions des systèmes d'information géographique (SIG) en archéologie est la modélisation, avec pour objectif de proposer une hypothèse spatiale plutôt que des données brutes. Cette contribution présente la conception, les résultats et les limites d'un SIG de modélisation spatiale appliqué à l'occupation rurale de la Somme à l'époque romaine. Il s'agit ici d'en délimiter les paramètres d'emplacement originels, c'est-à-dire les décisions anthropiques conditionnant la présence d'un site en un point précis. Dans la Somme, les variables influant statistiquement sur la distribution spatiale des fermes et *villae* relèvent plus notablement du domaine socio-économique que de l'environnement physique. Si l'aspect lacunaire de la documentation archéologique et des variables restituables est réel, il est indéniable que les SIG permettent d'atteindre une compréhension plus fine et plus fiable des paysages anciens et de leur interaction avec les habitants.

One of the functions of archaeological GIS is modelling, with the aim to propose a spatial hypothesis rather than raw data. This paper consists in presenting the design, results and limits of spatial modelling through GIS, which are applied to the rural settlement of the Somme in the Roman era. Its purpose is to ascertain locational parameters, *i.e.* the human decisions which condition the presence of a site in a specific area. In the Somme, variables that have a statistical impact on the spatial dis-

tribution of farms and *villae* tend to pertain more often to the socio-economic field rather than the physical environment. While the gap in the archaeological data and reconstituted variables is real, it is undeniable that GIS models allow a finer and more reliable understanding of ancient landscapes and their interaction with the inhabitants.

Entrées d'index

MOTS-CLÉS : humanités numériques spatialisées, analyse spatiale, archéologie, histoire, histoire antique, système d'information géographique, modélisation statistique

KEYWORDS: spatial digital humanities, spatial analysis, archaeology, history, ancient history, geographic information system, statistical modelling

Introduction

¹ La place fondamentale occupée par le monde rural au sein de l'Empire romain se justifie par son importance économique, démographique, culturelle et géographique. Si les estimations varient très largement d'une région à l'autre ou au cours du temps, la très grande majorité de la population de l'Empire, ainsi que de la plupart des autres sociétés antiques, faisait partie et dépendait du milieu rural pour sa subsistance et ses bénéfices. Afin d'appréhender le développement de l'économie rurale romaine, il apparaît nécessaire d'en comprendre les modalités de mise en place, c'est-à-dire le choix de localisation des établissements ruraux. Le rôle primordial pris par cette décision est ainsi relayé par les auteurs anciens, notamment les agronomes latins. Caton l'Ancien, Varron, Columelle puis Palladius mentionnent successivement divers facteurs environnementaux et socio-économiques à prendre en considération lors de la construction ou l'acquisition d'une propriété agraire (Nisard 1864).

² Malgré son importance relayée par la littérature ancienne, peu d'études concernant la Gaule du Nord prennent en considération les paramètres d'installation des établissements ruraux. Parmi les causes de cette lacune, les caractéristiques environnementales de cette partie de l'Empire expliquent un *a priori* assez péjoratif dans l'investigation de la question : la relative platitude topographique de cette partie de l'Empire serait la source d'une forte homogénéité de l'occupation, tout particulièrement dans la Somme, le département étudié dans le cadre de cette contribution (Ben Redjeb 2013, 188). Cette douceur topographique est en effet à l'origine du développement rapide et extensif de l'agriculture céréalière à champs ouverts dans la Somme, ce qui est matérialisé par la forte densité de sites ruraux déterminée par le biais des opérations archéologiques préventives, mais surtout des campagnes de prospections aériennes et pédestres réalisées entre les années 1960 et 1990. Ces dernières ont en effet permis la découverte d'environ 2 000 établissements ruraux dans le seul département de la Somme (Agache 1978 ; Bayard et Collart 1996 ; Ben Redjeb *et al.* 2005, 177-222 ; Ben Redjeb 2013 ; Bayard, Buchez et Depaepe 2014). Pour être valide, l'hypothèse d'une absence to-

tale de facteur d'influence – qu'il soit d'ordre naturel ou socio-économique – doit toutefois sortir du domaine empirique pour faire l'objet de constats mesurables et quantifiables, affirmant ou infirmant l'homogénéité de l'occupation. De plus, le relief ne saurait être le seul facteur d'influence dans la localisation des établissements ruraux. L'objectif de la présente étude est donc de déterminer quels paramètres peuvent influencer la localisation des établissements ruraux de la Somme durant la période impériale.

3 Afin de résoudre cette question avec un appui quantitatif, mesurable et reproductible, l'élaboration d'un système d'information géographique (SIG) constitue une étape indispensable. L'emploi de la plupart des SIG en archéologie fait partie d'une démarche illustrative plus qu'analytique, puisque l'illustration – la carte – constitue la finalité du processus, servant à donner une image de fond à des variables indépendantes (points et vecteurs géoréférencés). En revanche, la modélisation spatiale a une fonction analytique : la carte est ici une hypothèse, une résultante distincte de la réalité, qui est parfois même facultative lorsque l'objectif est uniquement de résoudre une problématique de quantification. Afin de traiter cette problématique archéologique, il conviendra de développer une approche plus méthodologique, permettant de débattre des avantages et limites des différentes approches d'analyse spatiale, y compris la modélisation prédictive. Mais premièrement, une définition plus précise du cas d'étude doit être réalisée.

Le cadre de l'étude

Caractéristiques géographiques

4 Le département actuel de la Somme constitue une aire d'étude dont les caractéristiques archéologiques et géographiques favorisent l'analyse de l'occupation rurale en conditions de quasi-homogénéité topographique. Les 6 170 km² de la Somme sont répartis entre des plateaux de basse altitude ainsi que des plaines alluviales ou maritimes. Il semblerait qu'une grande partie de la plaine maritime était submergée durant l'Antiquité, ainsi que l'indiquerait, d'une part l'absence de sites archéologiques romains, et, d'autre part l'analogie avec la région entre Canche et Authie, directement au nord dans le Pas-de-Calais. Dans cette région, l'influence maritime est dominante, hormis à l'emplacement des cordons dunaires encadrant les terres inondables, ces derniers pouvant constituer des îlots exondés et parfois occupés. Les limites du littoral ont beaucoup fluctué par le passé et ne sont pas encore suffisamment bien déterminées pour permettre une modélisation paléogéographique des plaines maritimes de la Somme, en particulier dans le Marquenterre, au nord-ouest de l'embouchure du fleuve Somme. À l'exception de ces plaines maritimes et en l'absence d'indices prouvant le contraire, la géographie physique de la Somme a peu ou pas évolué depuis l'Antiquité : elle peut donc être analysée par le biais des données topographiques actuelles.

5 S'il est tout à fait possible de réaliser une modélisation concernant l'intégralité du département, cela aurait pour conséquence d'atténuer considérablement les variations locales pouvant exister dans l'occupa-

tion d'un territoire de cette taille. On peut en effet noter de subtiles spécificités microrégionales : les plaines du sud du département – notamment le long de la rivière Selle – sont profondément incisées par des vallées encaissées, tandis que le Santerre à l'est du département est à l'inverse caractérisé par une topographie presque parfaitement homogène, le fleuve Somme n'y traçant qu'un mince sillon. De même, la géographie administrative antique ne concorde pas avec les divisions territoriales contemporaines, puisque le Santerre faisait partie de la cité des Viromanduels (capitale à Saint-Quentin puis Vermand durant l'Antiquité tardive), alors que la majorité de la Somme appartenait pour sa part à la cité des Ambiens (capitale à Amiens). Il n'est donc pas exclu que des choix différents aient pu déterminer la localisation des établissements ruraux dans différentes parties de la Somme.

6 L'approche microrégionale permet – par la subdivision en parties de dimensions comparables – d'inclure les variabilités locales dans l'interprétation des modèles. Trois critères ont sous-tendu la définition des quatre microrégions étudiées dans le cadre du présent article (figure 1) : la présence de caractères locaux marqués dans la géographie ancienne ou physique, un nombre de sites archéologiques suffisamment important pour faire l'objet de considérations statistiques (minimum 100) et enfin la qualité des données archéologiques recueillies, c'est-à-dire des résultats de fouilles préventives récentes venant appuyer les données issues de prospections pédestres et aériennes ou encore des découvertes fortuites et des fouilles anciennes. Toutes localisées le long du fleuve Somme et totalisant 2 090 km², les microrégions ont été déterminées spatialement par la création d'une zone tampon de 2 500 m autour des sites archéologiques sélectionnés, formant des formes globalement quadrangulaires de tailles analogues. Le tableau suivant fait la synthèse des composantes justifiant la définition de ces territoires (tableau 1).

FIGURE 1. CARTE DES QUATRE MICRORÉGIONS DE LA SOMME

Image produite par l'auteur, données archéologiques, © ABG, méthodologie © IHAPMA, MNT BD ALTI ®

TABLEAU 1. SPÉCIFICITÉS DES MICRORÉGIONS

	Spécificités locales	Nombre d'établissements ruraux	Nombre de sites fouillés (archéologie préventive ou sauvetage)	Pourcentage de sites fouillés par rapport au nombre total d'établissements ruraux
Micro-région 1	Basses terres peu encaissées par la Somme	208	17	8,17 %
Micro-région 2	Vallée très encaissée de la Selle se jetant dans la Somme, inclusion de la capitale ambiennaise à Amiens	171	35	20,47 %
Micro-région 3	Très nombreuses rivières et Somme méandreuse et encaissée	204	18	8,82 %
Micro-région 4	Topographie homogène de plateau et territoire faisant partie de la cité des Viromandueus	239	27	11,30 %

Cadre chronologique

7 Prendre en considération l'intégralité de la période romaine présente plusieurs biais importants. Premièrement, la période dite de « transition » entre la fin de la conquête césarienne en 50 avant Jésus-Christ et le règne d'Auguste présente des ruptures importantes, tant vis-à-vis de la fin de la période laténienne que des siècles suivants durant le Haut-Empire. De plus, les sites transitionnels – apparaissant et disparaissant à des rythmes difficiles à analyser – présentent une architecture en matériaux périssables qui, contrairement à la majorité des établissements ruraux du Haut-Empire, ne sont que très difficilement détectables autrement que par des opérations archéologiques préventives. L'organisation spatiale des établissements ruraux est parfois directement en continuité avec celle de la fin de la période laténienne, mais préfigure parfois celle du Haut-Empire, par le biais de « proto-*villae* » à l'instar de celle de Conchil-le-Temple. Néanmoins, des niveaux d'incendies ou de remaniements radicaux de l'organisation de ces fermes durant la période de transition pourraient être mis en lien avec les nombreux troubles et révoltes affectant les territoires voisins et probablement aussi celui des Ambiens et des Viromandueus. Les fouilles préventives n'ayant pas couvert la totalité des territoires compris au sein des micro-régions – bien que ces dernières soient parmi les zones les plus fouillées récemment – les établissements ruraux de transition ne sont pas encore assez nombreux pour offrir un portrait représentatif de l'occupation rurale à cette période. À l'inverse, le Haut-Empire est marqué par de nombreux établissements, documentés par les prospections aériennes profitant de la présence quasi systématique de soubassements en pierres, bien plus aisément détectables visuellement. Il apparaît donc impossible dans le cadre d'une modélisation et d'une analyse statistique de prendre en considération la période de transition. Le début de la période analysée se place donc au milieu du I^{er} siècle après Jésus-Christ.

8 La seconde difficulté chronologique concerne la fin de la période romaine. Les établissements ruraux occupés durant le Bas-Empire le sont presque exclusivement depuis le Haut-Empire, et les nouvelles créations

de sites sont donc particulièrement rares : seuls quatre établissements ruraux ont été édifiés après la première moitié du II^e siècle. Cette continuité d'emplacement – laquelle n'est pas nécessairement reflétée par une continuité architecturale – est totalement oblitérée entre la fin du III^e siècle et le début du V^e siècle, après quoi aucun établissement rural n'est connu jusqu'à la fin du V^e siècle voire du VI^e siècle. Si des facteurs de non-conservation des vestiges ou la superficie encore relativement faible des opérations archéologiques préventives pourraient expliquer une partie de ce hiatus, il n'est pas impossible que la dépopulation très importante du territoire soit la cause principale. Il est donc impossible de prendre en considération le V^e siècle dans cette étude. Les limites chronologiques sont donc placées au milieu du I^{er} siècle et à la fin du IV^e siècle.

- 9 Il convient désormais de faire la synthèse des données utilisées dans le cadre de la modélisation, ainsi que de la nature de l'occupation rurale au sein des quatre microrégions, conditionnant directement les choix d'implantation des établissements.

Données archéologiques et cartographiques

- 10 Le géoréférencement et la caractérisation des établissements ruraux de la Somme ont été effectués au sein du programme *ABG (Atlas des provinces romaines de Belgique et de Germanie)*, lequel est dirigé par Xavier Deru (université de Lille). Les quatre microrégions étudiées comportent un total de 822 établissements ruraux, dont 729 (89 %) sont des fermes bâties sur des soubassements en pierre, tandis que 61 autres (7 %) sont construits en matériaux périssables. Enfin, 32 sites possèdent une morphologie appartenant à celle de vastes *villae*. Cette classification en trois grandes familles d'établissements ruraux découle des efforts combinés de chercheurs français, belges et néerlandais (Ferdrière *et al.* 2010 ; Habermehl 2011 ; Deru 2012 ; Habermehl 2013). La *villa* y constitue une ferme bâtie sur soubassements de pierres, laquelle doit être pourvue de multiples bâtiments agricoles ainsi que d'un complexe résidentiel muni d'un minimum de 15 pièces, dont certaines peuvent posséder des systèmes de chauffage par hypocauste et parfois d'autres marqueurs de luxe (enduits peints, mosaïques, *opus sectile*, colonnes, etc.). Cette définition correspond à la catégorie de *villa* la plus développée spatialement et architecturalement. Mais elle est restrictive, les recherches anciennes sur la Gaule du Nord ayant eu tendance à attribuer le terme de *villa* à une large gamme de fermes comportant une ou deux cours alignées et des bâtiments s'y accolant.

- 11 Le processus de modélisation incluant des facteurs socio-économiques extérieurs au champ restreint de l'aire bâtie des établissements ruraux, 71 autres sites sont pris en considération. Il s'agit ici des capitales des cités des Ambiens et Viromanduens, mais aussi de sept capitales voisines ayant pu attirer l'implantation des établissements ruraux, au Haut-Empire ou au Bas-Empire. Il en va de même pour 62 agglomérations secondaires et sanctuaires, dont la présence à proximité des fermes a pu favoriser leur implantation et leur développement. Enfin, les routes romaines – dont le tracé a également été restitué au sein du programme *ABG* – forment des vecteurs et des carrefours ayant pu avoir eu une influence positive ou négative sur l'implantation rurale.

Indispensables dans le cadre d'une modélisation, les données spatiales primaires¹ comportent à la fois des fichiers vectoriels et *rasters*. Les vecteurs utilisés sont ceux des subdivisions administratives actuelles de la Somme ainsi que son réseau hydrographique, a priori relativement stable depuis l'Antiquité (données libres nationales). En l'absence de données précises sur la pédologie et la géologie de surface dans le secteur, aucune autre donnée environnementale vectorielle contemporaine ne peut être utilisée, car totalement anachronique (Corine Land Cover, couverture forestière de l'European Forest Data Center, etc.). La quasi-intégralité du processus de modélisation découle donc directement d'un seul jeu de données *raster* : le MNT (modèle numérique de terrain). Après plusieurs essais, il s'avère que l'échelle d'analyse du projet ne requiert pas un taux de précision par cellule de 1 ou 5 m, ces deux résolutions présentant une certaine « pollution » d'éléments topographiques contemporains (routes, plateformes construites, etc.). À l'inverse, la résolution de 75 m s'avère trop peu précise puisque les reliefs locaux sont particulièrement atténués. Une résolution intermédiaire de 25 m a donc été choisie, provenant du BD ALTI (©) distribué par l'IGN. Un postulat nécessaire à la démarche est que la topographie physique et l'utilisation générale des terres de la Somme ont peu changé depuis l'Antiquité, ce qui semble envisageable au vu des quelques esquisses sur le sujet (Wightman 1985, 4-5 ; Matterne 2001 ; Lepetz et Matterne 2003, 28).

L'ensemble des considérations préalables à la conception des modèles ayant été présenté, les avantages et inconvénients de plusieurs approches d'analyse spatiale vont désormais être débattus.

Méthodologie

Causalité et corrélation

14

La première difficulté rencontrée lors de tout processus d'analyse spatiale est la frontière très perméable entre les notions de causalité et de corrélation. C'est-à-dire qu'une variable qui semble être la cause de certains effets (dans notre cas une plus forte ou plus faible fréquence d'implantation rurale) pourrait plutôt être liée à une autre variable, voire à plusieurs, lesquelles peuvent avoir ou non un lien de causalité avec les résultats d'un modèle. Par analogie, la présente démarche modélisatrice peut être comparée à un sondage réalisé sur un échantillon de population (les établissements ruraux), et la relation de proportionnalité de telle ou telle variable appartenant à chaque individu (les paramètres de localisation) peut tout aussi bien être la conséquence d'une relation de cause à effet, que d'une corrélation avec d'autres variables non sondées, c'est-à-dire d'une relation proportionnelle entre deux ou de multiples valeurs numériques. L'exemple emblématique de ce problème est celui rapporté par Harris et Lock (2006, 49) : la variable semblant influencer le plus l'implantation humaine précolombienne en Virginie occidentale est la présence du platane d'Amérique sur le territoire actuel. Il est évident que, dans ce cas, tant la présence de platanes que l'implantation humaine ancienne sont à mettre en relation avec une combinaison d'autres variables non modélisées les conditionnant de façon similaire (la présence d'eau à faible profondeur par exemple).

15

Cette conséquence dichotomique du processus de modélisation est inexorable et insoluble. Comme noté par Verhagen et Whitley (2012, 57), même avec un support statistique, nulle relation causale ne peut être prouvée. De ce fait, la production de modèles et leur analyse statistique constituent uniquement un moyen de mesurer et « matérialiser » des hypothèses spatiales, lesquelles ne peuvent en aucun cas être vérifiées, puisque les faits étudiés ne peuvent être reproduits. En revanche, en l'absence de modèles mesurables, toute théorie sur les facteurs d'implantation humaine ne pourrait sortir du cadre d'une argumentation circulaire : les paramètres de cet emplacement sont favorables puisqu'un établissement rural s'y trouve et un établissement rural s'y trouve puisque l'emplacement est favorable. L'emploi de la modélisation dans une démarche explicative, s'il doit être modéré, reste néanmoins indispensable. Ce constat étant fait, la réalisation et l'analyse de modèles spatiaux s'insère dans une sous-discipline souvent décriée : la modélisation prédictive.

Les modélisations prédictives

16 À cheval entre l'écologie, les sciences informatiques et les sciences humaines, la modélisation prédictive prend diverses formes et utilisations selon les disciplines mères. Une définition générale de la modélisation prédictive est celle de « prédire la localisation de sites archéologiques ou de matériel dans une région, sur la base d'un échantillon de cette région ou de notions fondamentales concernant les comportements humains » (Kohler et Parker 1986, 400). Dans cette définition sont contenues les deux familles de protocoles de modélisation : celle fondée sur les variables dépendantes (les sites) et celle fondée sur les variables indépendantes (des classes d'espace).

17 Cette distinction n'est pas triviale, car lorsque la modélisation est basée sur les sites, son but est de prédire l'emplacement de sites inconnus, ce qui est particulièrement utile dans le cas de la gestion du patrimoine archéologique, par exemple afin de produire des cartes de sensibilité archéologique. À l'inverse, en fondant la démarche sur les classes d'espace, on peut mieux comprendre les paramètres d'implantation des sites déjà connus (Wheatley 2004, 5 ; Whitley 2004a, 2). Cette dernière démarche est tout aussi valide dans le cas de l'analyse spatiale archéologique que dans la biologie, les espèces animales ou végétales pouvant dans ce cas constituer l'échantillon devant être étudié. Dans le premier protocole, les variables indépendantes sont construites afin de corrélérer avec la distribution de sites existants (*data-driven modelling*), tandis que dans le second cas, ces variables indépendantes sont manuellement conçues par l'opérateur du SIG sur la base d'hypothèses spatiales (*theory-driven modelling*), les sites existants n'y intervenant que pour fournir un rapport statistique permettant ou non d'évoquer une corrélation entre les modèles fictifs et les données réelles. Si l'objectif d'une modélisation se rapporte nécessairement à l'une ou l'autre méthode, une certaine perméabilité méthodologique permet désormais d'employer les deux approches sans exclusion mutuelle (Verhagen et Whitley 2012, 90-91). Toutefois, la présente étude appartient essentiellement à la modélisation fondée sur les hypothèses spatiales.

18 Les premiers travaux de modélisation prédictive ont été publiés dès les années 1970 au sein des deux démarches fondamentales susmentionnées (Kvamme 2006). En ce qui concerne l'analyse des paramètres de localisation de sites archéologiques, l'étude réalisée par Jochim (1976) constitue à la fois la première tentative dans ce domaine et un exemple emblématique et exhaustif de cette démarche, malgré l'absence de SIG. Pour chaque cellule de l'espace défini dans cette étude, un vaste ensemble de propriétés environnementales est calculé (potentiel calorifique, démographique, écologique, etc.), puis les cartes sont dessinées manuellement selon ces critères numériques, et enfin elles sont confrontées à la répartition réelle des indices de passage des sociétés de chasseurs-cueilleurs. Le développement et la démocratisation des ordinateurs et des SIG n'ont nullement bouleversé cette façon de procéder, mais ont simplement facilité l'exécution des étapes intermédiaires entre la conception d'hypothèses sur les comportements humains et l'évaluation de leur interaction statistique avec les données réelles. On retrouve déjà les variables environnementales désormais archétypiques de cette

démarche : « les pentes, l'altitude, l'orientation des sols, le relief local, les crêtes, les lignes de drainage, la variation de terrain et les surfaces de distance pour les vecteurs de cours d'eau » (Kvamme 2006, 10). Les premiers résultats de modélisation prédictive sur SIG sont produits en 1985 (Judge et Sebastian 1988), après quoi une certaine réticence à l'emploi et au développement de ces logiciels s'installe en archéologie, tout particulièrement dans le milieu des chercheurs anglo-saxons. En effet, leurs critiques importantes – et souvent justifiées – amenèrent l'analyse spatiale, surtout concentrée sur des critères environnementaux, à subir un quasi-anathème visant le caractère de « fausseté écologique ou de déterminisme environnemental » de la démarche (Wheatley 2004, 6-7).

La modélisation de facteurs sociaux, culturels et économiques

19 En réponse et en réaction aux critiques post-processuelles de la modélisation prédictive, les vingt dernières années ont vu le développement d'approches théoriques et de cas d'étude concernés par des variables spatiales ne dépendant généralement pas de l'environnement physique (Kvamme 2006, 21). On peut citer les travaux précoces analysant l'impact de la visibilité des environs sur la localisation de fortification croates (Stančić et Kvamme 1999), ou encore le programme *ArchaeDyn* au sein duquel plusieurs chercheurs incorporèrent d'autres types d'analyses de la visibilité, de la proximité et de caractères intrinsèques de sites archéologiques dans de multiples études de cas (Gandini, Favory et Nuninger 2012). Ce sont en particulier les collaborations franco-néerlandaises qui fournissent des nouveautés dans l'exploitation de variables non environnementales au sein du projet *IHAPMA* (*Introducing the Human (f)Actor in Predictive Modelling for Archaeology*) auquel nous avons contribué. Ainsi, on note l'addition de variables incluant des calculs de distance-coût (chemins de moindre coût) et de visibilité cumulée (la visibilité depuis et vers un large ensemble de points), ou encore de la persistance de l'anthropisation des paysages (Verhagen *et al.* 2013, 2015 ; Nüsslein 2016 ; Nuninger *et al.* 2016).

20 Au cours de la présente étude, largement inspirée des derniers travaux du projet *IHAPMA*, l'accent porté sur les variables socioculturelles et économiques permet d'évoquer de nouvelles variables, notamment la distance-coût multivariée, la prééminence des sites élitaires, mais surtout, cette étude vise à comparer les influences relatives ou combinées des variables environnementales, socioculturelles et économiques.

Modèles et résultats

Exemple détaillé de la démarche

- 21 Afin de rendre plus intelligible la démarche, la réalisation et l'évaluation de la première variable, dans le cas présent la pente du terrain, va être détaillée intégralement (figure 2).

FIGURE 2. MODÈLE RECLASSÉ DES PENTES

La catégorie en rouge est supposée être la plus prédictive, tandis que la jaune représente la classe moyenne et la verte la classe la moins propice à l'implantation d'établissements ruraux.

Image produite par l'auteur, données archéologiques © ABG, méthodologie © IHAPMA, MNT BD ALTI ®

- 22 Pour chaque variable, le processus de modélisation débute par la formulation d'une théorie, ici celle que les fermiers romains ont pu choisir de préférence des sites d'installation comportant très peu de pente, favorisant les cultures céréalières extensives caractéristiques de la période.
- 23 Afin de confirmer ou d'infirmer cette théorie, elle doit d'abord être traduite au sein du SIG par le biais de la reclassification du modèle de pentes, lequel dérive du MNT. En effet, ce dernier, lorsqu'il est produit par un SIG, n'est dans un premier temps pas classé, mais présente des valeurs continues (de 0 à 100 %). La classification choisie pour cette variable peut être appelée « arbitraire » puisqu'elle est déterminée manuellement, selon des critères empiriques de pénibilité exponentielle de déplacement selon la pente. Ainsi, les régions présentant des pentes supérieures à 6 % sont pressenties comme particulièrement néfastes à l'installation humaine, et l'ensemble des cellules présentant des valeurs égales ou supérieures à 6 % sont alors converties en une seule valeur dite « prédictive », celle de 1. Les pentes comprises entre 6 % et 2 % sont pres-

senties comme moyennement favorables à l'installation humaine, permettant un bon écoulement des eaux mais un déplacement assez entravé, leur permettant de se voir attribuer une valeur prédictive médiane de 2. Enfin, les cellules dont le pourcentage de pente est inférieur à 2 % se voient attribuer une valeur prédictive haute de 3. La simplicité apparente de cette méthode de classification est voulue par la nécessité de conserver un protocole fondé intégralement sur les hypothèses comportementales. Les mêmes classes sont établies dans les quatre microrégions, différant nettement en matière de représentation. Ainsi, alors que les régions 2 et 3 comportent 15 % de territoires à fortes pentes, la région 4 en est quasi dépourvue (0,03 %). D'autres méthodes de classification, par exemple les quantiles (chaque classe présente autant de cellules que les autres), auraient pour effet de supprimer toute logique comportementale à la démarche, puisque classer les pentes de la sorte entraîne la création de trois classes peu distinctes car très rapprochées les unes des autres en ce qui concerne le pourcentage de pente.

24 Une fois la reclassification du modèle univarié réalisée, celui-ci peut être confronté à la répartition réelle des établissements ruraux. La procédure, semi-automatisée par l'intermédiaire du *model builder* d'ArcGIS, consiste tout d'abord, microrégion par microrégion, en un calcul du pourcentage des cellules comportant chaque valeur prédictive (P_a , de 0 à 1), puis du pourcentage de sites présents dans chaque classe (P_o , également de 0 à 1). Ces deux valeurs seules peuvent être utilisées afin de déterminer les valeurs prédictives de chaque classe. La modélisation prédictive telle que pratiquée le plus fréquemment emploie jusqu'à quatre valeurs prédictives différentes : le gain de Kvamme (P_a/P_o), la valeur indicative (P_o/P_a), le gain relatif ($P_o - P_a$) et le test khi-deux (probabilité que les deux ratios soient indépendants, et donc non corrélés, dans le cas où la valeur obtenue est supérieure à 0,05).

25 Tous les modèles univariés et multivariés réalisés dans la Somme ne présentant pas d'amplitudes prédictives extrêmes, seuls la valeur indicative et le khi-deux ont été analysés (tableau 2). Dans l'exemple de la pente, la valeur indicative des classes déterminées varie de 0,39 à 1,19. En d'autres termes, on peut en certains endroits trouver 39 % du nombre de sites attendus si la pente n'avait aucune influence, tandis que d'autres classes présentent jusqu'à 119 % du nombre de sites attendus. Dans les trois premières régions, les classes déterminées présentent une corrélation assez haute avec les établissements ruraux, c'est-à-dire que la fréquence d'établissements ruraux est plus importante dans les régions de basse pente par rapport aux régions de moyenne pente, également plus attractives que les régions de forte pente. La microrégion 4 dévie de ce constat positif, mais la très faible représentation des zones de moyennes et hautes pentes (8 % à elles deux) implique une très forte homogénéité topographique. Les tests de khi-deux sont en concordance avec les valeurs indicatives, puisqu'ils présentent des valeurs très basses dans les trois premières régions, tandis que la quatrième, avec une valeur nettement supérieure au seuil de 0,05 %, présente une indépendance forte entre le pourcentage de sites et le pourcentage de surface.

TABLEAU 2. RÉSULTATS STATISTIQUES DU MODÈLE UNIVARIÉ DE PENTES

Classe de pente	Classe prédictive	Sites réels	Sites attendus en état d'homogénéité	Pourcentage de la superficie P_a	Pourcentage des sites P_o	Valeur indicative (P_o/P_a)	Test khi-deux
Région 1							
6-82 %	1	15	30,58	0,15	0,07	0,49	0,00033
2-6 %	2	18	28,15	0,13	0,09	0,64	
0-2 %	3	175	149,28	0,72	0,84	1,17	
Région 2							
6-82 %	1	10	25,52	0,15	0,058	0,39	0,00029
2-6 %	2	16	23,56	0,14	0,094	0,68	
0-2 %	3	145	121,92	0,71	0,85	1,19	
Région 3							
6-82 %	1	10	21,83	0,11	0,05	0,46	0,00249
2-6 %	2	18	26,74	0,13	0,09	0,67	
0-2 %	3	176	155,43	0,76	0,86	1,13	
Région 4							
6-82 %	1	5	6,30	0,03	0,02	0,79	0,23481
2-6 %	2	7	12,50	0,05	0,03	0,56	
0-2 %	3	227	220,20	0,92	0,95	1,03	

26

À l'issue de ces tests statistiques particulièrement simples, les résultats du modèle peuvent être confrontés à l'hypothèse de départ. Dans le cas des pentes, on peut affirmer que les terrains présentant de fortes pentes sont nettement défavorisés, sans toutefois être proscrits, et il en est de même pour les pentes moyennes. Les zones de basse pente, si elles sont toujours plus attractives que les autres, ne présentent néanmoins pas de prédictivité suffisamment haute pour constituer une variable fondamentale dans le processus de décision de la création d'un établissement rural.

27

L'intégralité de la démarche sus-développée peut alors être répliquée à l'ensemble des variables dépendantes, mais aussi à toutes les subdivisions des variables indépendantes (type de site, période d'occupation, etc.), ce qui représente un nombre considérable de modèles à réaliser et évaluer. La procédure est donc très chronophage malgré la semi-automatisation mise en place.

Les modèles univariés et leurs hypothèses

28

Dans cette partie, nous résumons les caractéristiques, méthodes de classification et hypothèses permettant l'analyse de douze variables différentes.

Outre la pente déjà présentée, les variables environnementales incluent l'orientation des sols, leur ensoleillement, leur classe topographique ainsi que la distance-coût par rapport aux cours d'eau. L'orientation des sols, fondamentale dans l'agriculture méditerranéenne telle que décrite par les agronomes latins, pourrait encore avoir une certaine influence sur l'agriculture du nord de la Gaule. Elle est calculée par le SIG sous la forme de valeurs allant de -1° (terrain plat) à 360° d'angle. Leur reclassification est facilitée par la détermination des huit directions cardinales, ensuite réduite à trois classes supposées refléter des caractères plus ou moins attractifs (ou prédictifs) des sols d'installation : le sud et le sud-ouest en classe hautement attractive ; l'ouest, l'est, le sud-est et les zones plates en classe moyennement attractive ; et enfin le nord, le nord-ouest et le nord-est en classe peu prédictive. L'ensoleillement, corrélé et calculé par le biais de l'orientation des sols, dépend également de l'environnement voisin qui pourrait porter ombrage à certains moments de la journée. Un fort ensoleillement pourrait là aussi favoriser une agriculture céréalière, et les valeurs continues d'ensoleillement, reclassées en trois quantiles, présentent des valeurs allant de $3\,223\text{ Wh/m}^2$ à $5\,658\text{ Wh/m}^2$, par ordre de prédictivité supposée.

La classe topographique des sols, c'est-à-dire le type de milieu physique, allant de la vallée encaissée aux plaines et crêtes, peut être modélisée par différents algorithmes, notamment le TPI (*topographic position index*) et le DEV (*derivation from mean elevation*). Le DEV s'avère un algorithme plus adéquat envers l'environnement de la Somme, puisqu'il permet de distinguer des reliefs peu accentués, alors que le TPI atténue ces derniers. La préclassification du DEV par écart type permet de distinguer sept classes topographiques, reclassables en trois catégories de prédictivité a priori en lien avec une agriculture à champs ouverts : les vallées encaissées et autres dépressions profondes constituent la classe supposée la moins attractive pour l'installation d'établissements ruraux, tandis que les bas de versants, les vallées peu encaissées et les sommets de versants constitueraient une classe moyennement attractive, supplantée par les plaines, plateaux et versants supérieurs qui devraient offrir les meilleures conditions pour l'agriculture et la construction.

La dernière variable environnementale employée, la distance-coût aux cours d'eau, modèle l'accessibilité de voies navigables ou exploitables pour la pêche ou d'autres activités annexes de la vie agricole. À l'inverse, une trop grande proximité avec les cours d'eau – ainsi que l'ont noté les agronomes latins – pourrait entraîner des inconvénients importants : inondations, sols trop humides ou encore de graves maladies. Généralement employée afin de modéliser l'accessibilité d'un ou plusieurs points à d'autres, l'analyse de chemins de moindre coût peut également être utilisée avec des vecteurs, ici les cours d'eau. Plutôt qu'une distance euclidienne ne prenant pas en compte les obstacles topographiques, cette approche permet de simuler un parcours anthropique plus crédible, dépendant de la pente par l'intermédiaire de l'algorithme de Tobler (Tobler 1993 ; Herzog 2013 ; Lock et Pouncett 2010 ; Verhagen et Jeneson 2012 ; Güimil-Fariña et Parcero-Oubiña 2015). Le résultat, exponentiellement proportionnel à la pente, est exprimé en heures de voyage. La classe estimée la moins prédictive regroupe les cellules localisées à moins d'une minute de trajet d'un cours d'eau. Entre une minute et une demi-heure de trajet ainsi qu'au-delà d'une heure de trajet, la distance aux cours

d'eau pourrait avoir une influence moyenne sur l'implantation des établissements ruraux, tandis que les régions éloignées d'une demi-heure à une heure d'un cours d'eau pourraient en bénéficier le plus, étant suffisamment éloignées de ce dernier pour ne pas en subir les inconvénients et demeurant assez près pour y accéder en très peu de temps.

32

En ce qui concerne les variables socioculturelles et environnementales, elles dépendent de deux types d'analyse : l'analyse de distance et celle de la visibilité. Les modèles de proximité aux capitales de cité et aux marchés locaux (agglomérations secondaires et sanctuaires) constituent des distances-coût analogues à celle mise en place dans le cas des cours d'eau ; leur reclassification diffère toutefois très largement. Quant aux routes romaines, leur influence pourrait être moindre au-delà de plusieurs kilomètres, et une modélisation euclidienne semble suffisante. Toutes suivent la même hypothèse : plus grande est la proximité à ces trois catégories de peuplement et d'aménagement, plus grande devrait être la fréquence des sites, selon un postulat d'attractivité sociale, culturelle et surtout économique. La reclassification de ces variables est détaillée dans le tableau ci-dessous (tableau 3). On notera que la classification de la distance aux capitales de cité est plus permissive, ce type de site pouvant indéniablement constituer un point d'intérêt fondamental dans la décision d'implantation d'un établissement rural.

TABLEAU 3. CLASSIFICATION DES VARIABLES SOCIO-ÉCONOMIQUES

	Valeur prédictive	Classification
Distance euclidienne aux routes (en mètres)	1	2 000-97 870 m
	2	1 000-2 000 m
	3	0-1 000 m
Distance-coût aux capitales de cité (en heures)	1	2,25-10 h
	2	1,75-2,25 h
	3	0-1,25 h
Distance-coût aux marchés locaux, soit les agglomérations secondaires et sanctuaires (en heures)	1	1-10 h
	2	0,5-1 h
	3	0-0,5 h

33

Deux autres variables sont modélisées par le biais de distances euclidiennes : l'attraction d'établissements ruraux par les grandes *villae* et la proximité de la frontière. La première matérialise l'hypothèse selon laquelle la préexistence d'une *villa* à proximité d'un emplacement favoriserait l'implantation d'un nouvel établissement (en fondations de pierre ou en matériaux périssables). Dans ce cas, il est supposé qu'être à moins de 1 000 m d'une *villa* serait le meilleur cas possible, tandis que la classe moyenne se situe entre 2 000 et 5 000 m, et la classe supposée défavorable au-delà de 5 000 m. À l'inverse, le modèle de proximité à la frontière – ici celle des Ambiens et des Viromanduens, qui peut être étudié uniquement dans les microrégions 3 et 4 – serait plutôt un vecteur négatif de l'implantation, les fortins s'y situant requérant des ressources prélevées dans les environs. Si, au contraire, on considère l'hypothèse tout aussi valide voyant la frontière comme un vecteur bénéfique d'échanges

entre les territoires voisins, les résultats du modèle devraient être inversés. Pour cette variable, les seuils de classe sont également placés à 2 000 et 5 000 m.

34

Les deux dernières variables sont en rapport avec les analyses de visibilité. La topographie largement ouverte de la Somme implique une visibilité particulièrement haute depuis presque tous les emplacements de l'espace. Toutefois, certains terrains pourraient se trouver au-dessus des environs voisins, leur conférant une qualité de « proéminence » (figure 3), laquelle – si elle est recherchée par les propriétaires d'établissements ruraux – pourrait offrir deux avantages : la surveillance aisée sur les dangers alentour (brigandage, armées), ainsi que l'acquisition d'un panorama offrant une certaine esthétique d'emplacement et de mise en valeur de la ferme. L'algorithme employé est celui dit d'ouverture (*openness*), développé et distribué par OPALS (université de Vienne). Ce dernier permet de calculer des azimuts positifs et négatifs, c'est-à-dire des angles de visions vers le haut mais aussi vers le bas, ces derniers étant dans notre cas caractéristiques d'une proéminence. La reclassification de ce modèle est réalisée par écart type.

FIGURE 3. MODÈLE UNIVARIÉ DE LA PROÉMINENCE

La classe d'espace en rouge indique des emplacements ayant une dominance visuelle sur leurs environs.

Image produite par l'auteur, données archéologiques © ABG, méthodologie © IHAPMA, MNT BD ALTI®

35

Pour finir, la dernière variable modélisée concerne la visibilité des sanctuaires, c'est-à-dire la présence d'un sanctuaire dans le champ de vision d'un potentiel établissement rural (figure 4). Hautement hypothétique, ce modèle verrait les propriétaires d'établissements ruraux choisir des emplacements ayant vue sur un sanctuaire, afin d'être sous sa protection. L'analyse cumulative de visibilité prend ici sa source dans les sanctuaires, avec un décalage altimétrique de 5 m, afin de surélever légèrement les constructions désormais arasées de ces édifices. Il s'agit ici de la seule variable n'ayant pas été reclassée en trois catégories, mais

plutôt en deux valeurs binaires : visibilité ou non-visibilité. Toutes les variables ayant été présentées, leur influence doit désormais être évaluée.

FIGURE 4. MODÈLE UNIVARIÉ DE LA VISIBILITÉ DES SANCTUAIRES

La classe d'espace en rouge indique les emplacements depuis lesquels un sanctuaire est potentiellement visible, tandis qu'aucun ne peut être aperçu depuis la classe d'espace en jaune.

Image produite par l'auteur, données archéologiques © ABG, méthodologie © IHAPMA, MNT BD ALTI®

Résultats statistiques et modèles multivariés

36

Le tableau ci-dessous (tableau 4) résume de façon simplifiée les résultats de l'analyse des modèles univariés, y indiquant les minimales, moyennes et maximales de la valeur indicative prise sur l'ensemble des microrégions. L'exactitude caractérise le degré de concordance entre la valeur indicative d'une classe et sa valeur prédictive au sein de l'hypothèse de départ. Le degré d'amplitude dénote la variation maximale de valeurs indicatives au sein du modèle, ici synonyme d'une influence positive et/ou négative forte. Le degré de cohérence est influencé par le nombre de microrégions suivant avec exactitude l'hypothèse de départ, mais aussi et surtout par le résultat du test de khi-deux. Enfin, la validité est un résumé empirique des trois caractéristiques précédentes, fournissant ou non l'indice d'une variable à forte influence potentielle sur l'implantation des établissements ruraux.

TABLEAU 4. RÉSULTATS STATISTIQUES SYNTHÉTISÉS DE TOUS LES MODÈLES UNIVARIÉS

	Min.	Moy	Max.	Exacti- tude	Ampli- tude	Cohé- rence	Validité
Pente	0,39	0,64	1,19	haute	moyenne	moyenne	haute
Orientation des sols	0,82	1,10	1,00	basse	basse	basse	basse
Ensoleillement	0,73	1,23	0,94	basse	basse	basse	basse
Classe topographique	0,38	0,74	1,25	haute	moyenne	moyenne- haute	haute
Distance-coût aux rivières	0,21	0,85	1,26	moyenne	moyenne	moyenne	moyenne
Distance euclidienne aux routes	0,41	1,12	1,65	haute	haute	moyenne- haute	haute
Distance-coût aux capitales de cité	0,25	0,94	1,26	haute	moyenne- haute	moyenne- haute	haute
Distance-coût aux marchés locaux (agglomé- rations secon- daires et sanctuaires)	0,26	0,92	1,37	haute	moyenne- haute	haute	très haute
Distance euclidienne aux <i>villae</i>	0,94	1,02	0,97	basse	basse	moyenne	basse
Proéminence	1,17	0,97	0,84	basse	basse	moyenne	basse
Visibilité des sanctuaires	0,96	1,10	1,22	moyenne	basse	basse	basse
Distance euclidienne aux frontières	1,04	1,13	0,72	basse	basse	basse	basse

37

D'un point de vue environnemental, peu d'emplacements sont réellement restrictifs envers l'occupation rurale, puisque les classes prédictives minimales – même si elles représentent parfois une part infime de la surface totale des microrégions – sont tout de même pourvues de sites archéologiques. C'est tout particulièrement le cas dans la quatrième région, celle au relief le plus doux et présentant fréquemment un caractère homogène relayé par les résultats de ses tests de khi-deux. Toutefois, on peut affirmer que certaines classes sont quantitativement plus attrayantes : la présence d'une rivière entre une demi-heure et une heure de marche, de faibles pentes et une position sur les plaines, les plateaux ou les versants supérieurs. Un deuxième constat important à évoquer est l'absence totale de corrélation entre la présence d'établissements ruraux et les modalités d'orientation et d'ensoleillement des terrains, présentant des résultats statistiques aléatoires et des résultats de test khi-deux toujours au-delà du seuil sous lequel ces variables pourraient être proportionnelles à l'occupation humaine.

38

En ce qui concerne les variables culturelles, sociales et économiques, on peut également éliminer la proximité des *villae*, la proéminence, la visibilité des sanctuaires et la distance à la frontière, clairement sans lien avec l'installation des fermes romaines, en tout cas dans les micro-

régions étudiées et dans la période concernée. Sans trop de surprise, la proximité des routes, des capitales de cité et surtout des marchés locaux constitue un facteur attractif. Les valeurs de prédictivité les plus hautes et les plus basses sont d'ailleurs atteintes pour ces trois variables, présentant une forte cohérence à travers les quatre régions, en particulier en ce qui concerne la présence de marchés locaux, puisque la quatrième région, généralement indifférente à tout paramètre, y est également sensible. Néanmoins, la causalité entre ces facteurs et l'installation des établissements ruraux pourrait, au moins en partie, être inversée, c'est-à-dire que la présence de fermes pourrait favoriser l'implantation d'agglomérations, de voies et de marchés.

39

Six variables constituant des paramètres apparemment corrélés avec l'occupation rurale romaine de la Somme, il est désormais possible de les combiner afin de vérifier leur compatibilité, voire leur effet additif sur les choix d'installation. L'hypothèse de base, finalement très peu distincte de ce qui est généralement formulé empiriquement dans les études monographiques de l'occupation rurale du nord de la Gaule, consiste en la potentielle supériorité d'influence des critères socio-économiques sur les critères environnementaux. Pour tester cela, il convient tout d'abord de combiner les variables environnementales, puis ces dernières avec les trois autres facteurs, lesquels devraient – afin de valider l'hypothèse – conférer des valeurs prédictives plus amples et cohérentes. La procédure utilisée pour la création de ces modèles multivariés est très simple : l'algèbre spatiale de la calculatrice *raster*. Il s'agit de définir une formule additive ou multiplicatrice de plusieurs variables spatiales dont voici le décomposé pour le modèle multivarié environnemental :

$$((DEV \times 1,5) + (pentes \times 1) + (distance \text{ aux cours d'eau} \times 0,5)) / 3$$

40

Les facteurs multiplicateurs ne sont pas totalement arbitraires, mais suivent à la fois les indications des modèles univariés et l'hypothèse empirique d'une plus grande importance du type de milieu topographique par rapport aux autres facteurs. Les agronomes latins n'ayant pas classé par ordre d'importance les facteurs décisionnels d'implantation rurale, il n'est pas possible de s'inspirer de ces derniers afin de pondérer les paramètres spatiaux. Une fois le modèle reclassé par écarts types (les modèles le composant étant déjà classés de 1 à 3), on obtient donc le modèle présenté ci-dessus (figure 5). Il peut ensuite suivre la même procédure d'évaluation statistique que celle précédemment développée. Pour ce modèle multivarié, la valeur prédictive maximale atteinte dépasse celle des modèles univariés le composant et il en est de même de l'exactitude, de l'amplitude, de la cohérence et de la validité (tableau 5). On peut donc affirmer que les paramètres ainsi pondérés ont un effet additif permettant une meilleure reconstitution des choix effectués par les propriétaires des établissements ruraux vis-à-vis de leur environnement physique.

FIGURE 5. MODÈLE MULTIVARIÉ DE L'ENVIRONNEMENT PHYSIQUE (CLASSE TOPOGRAPHIQUE, PENTES ET PROXIMITÉ DES COURS D'EAU)

Image produite par l'auteur, données archéologiques © ABG, méthodologie © IHAPMA, MNT BD ALTI ®

TABLEAU 5. RÉSULTATS STATISTIQUES SIMPLIFIÉS DES MODÈLES MULTIVARIÉS

	Min.	Moy	Max.	Exacti- tude	Ampli- tude	Cohérence	Validité
Modèle multivarié environnemental	0,46	0,65	1,41	haute	haute	moyenne- haute	haute
Modèle multivarié environnemental et socio-économique	0,11	1,02	1,52	très haute	haute	haute	très haute

41 La combinaison de ce modèle multivarié environnemental avec les variables socio-économiques suit la formule suivante :
((modèle multivarié environnemental × 1,6) + (distance aux marchés locaux × 1) + distance aux capitales de cité × 0,8) + (distance aux routes × 0,6)) / 4

42 On notera dans ce nouveau modèle visible ci-dessous (figure 6) qu'il a été délibérément choisi d'alourdir le poids des facteurs socio-économiques par rapport au modèle multivarié, les premiers totalisant 2,4 parts sur les 4 disponibles.

Image produite par l'auteur, données archéologiques © ABG, méthodologie © IHAPMA, MNT BD ALTI®

43 Les valeurs statistiques obtenues pour le modèle final sont les plus convaincantes (tableau 5), avec une amplitude très forte, une classe basse portant une prédictivité très faible et une classe haute avec une prédictivité haute. Toutes les microrégions sauf la quatrième réagissent similairement au modèle réalisé, qui semble en tout état de fait restituer là aussi un processus additif dans l'apport des variables socio-économiques, améliorant considérablement les résultats statistiques. La nature des variables conditionnant l'installation rurale dans la quatrième microrégion nous échappe et devrait être recherchée ailleurs, voire être totalement absente ; il n'est pas impossible que l'occupation soit quasi homogène et également répartie dans ce territoire.

44 Enfin, tous les modèles précédemment décrits se comportent de façon identique durant la première moitié du II^e siècle ou la première moitié du IV^e siècle, ce qui est également le cas pour les trois types d'établissements ruraux, hormis une plus forte attraction des variables socio-économiques pour la construction des *villae*, ce qui appuie l'hypothèse selon laquelle ces établissements constituent des éléments interconnectés d'un réseau plus vaste, plutôt que des noyaux autarciques. L'influence des variables socio-économiques sur les établissements plus modestes plaide aussi en faveur de l'hypothèse de fermes à l'économie dynamique plutôt que de celle de petites dépendances isolées dans les secteurs marginaux des territoires (Blondiau 2014, 7-11).

Conclusions

45

Par cette approche de modélisation supportée par des hypothèses spatiales, nous espérons avoir présenté une démarche aisément répliquable et simple d'accès grâce à la transparence de ses fondements et la liberté qu'elle donne dans la reclassification de ses modèles et l'élaboration de nouvelles théories. L'importance fondamentale des paramètres sociaux, économiques et culturels dans l'implantation humaine a pu être évaluée et semble en effet prendre une grande part dans les décisions anthropiques de cette aire géographique spécifique. Les résultats de cette étude montrent également le rôle non déterministe des variables mises en œuvre, non seulement dans la présence d'établissements ruraux, mais également dans leur absence. Aucune catégorie d'espace n'est dépourvue de sites et les catégories en présentant la plus grande fréquence ne sont pas attractives au-delà de 1,65 fois la normale. Si la forte homogénéité environnementale ou d'occupation pourrait être mise en cause dans ce constat, indiquons également que de nombreuses variables n'ont pas pu être modélisées, lesquelles sont pourtant citées par les agronomes latins : la qualité agraire de la terre, la salubrité de l'air et de l'eau, du climat, des forêts environnantes, la qualité de l'aide fournie par les voisins d'une propriété agricole ou encore la dangerosité des campagnes environnantes.

46

Non seulement les données archéologiques, pédologiques et palynologiques manquent pour reconstituer cet environnement passé, mais la présence de brigands et la qualité des voisins ne peuvent en aucun cas être quantifiées. L'apparente infinité de variables pouvant être modélisées et évaluées doit donc être largement modérée à ce qu'il est archéologiquement possible d'affirmer sur les espaces antiques. Plus encore que les biais propres à la démarche modélisatrice, c'est la donnée archéologique qui doit être critiquée : tous les terroirs ont-ils été prospectés de façon uniforme ? Analysons-nous les espaces favorisant la fondation d'établissements ruraux ou simplement leur conservation partielle dans les sous-sols ? Enfin et évidemment, comprendre l'occupation gallo-romaine nécessiterait de comprendre aussi bien celle qui la précède immédiatement. Or ce n'est pas le cas, tant quantitativement que qualitativement, puisque seules les opérations d'archéologie préventive peuvent documenter de manière fiable les occupations protohistoriques. D'ailleurs, seules les données issues de fouilles pourraient fournir suffisamment d'informations pour permettre la modélisation diachronique par état de construction, et donc l'évolution des établissements ruraux au cours du temps et en fonction des paramètres spatiaux les conditionnant. Les problèmes souvent insolubles de cette démarche modélisatrice ont été soulevés maintes fois, non seulement par ses détracteurs mais aussi par ses pratiquants, notamment Hans Kamermans (2004, 273).

47

Malgré tout, si chaque individu du passé a dû avoir un processus décisionnel différent et spécifique (Whitley 2004b, 4), certaines tendances globales partagées par la plupart des acteurs peuvent être mises en exergue, y compris dans un environnement aussi peu contraignant que celui de la Somme (Gaffney et van Leusen 1995, 371). Rappelons que l'apparente modestie statistique des résultats des modèles présentés n'em-

pêche en rien d'apporter un soutien important à des hypothèses qui étaient dépourvues d'appui quantitatif. Des modèles fondés sur la répartition des sites auraient évidemment une performance statistique supérieure, mais ne présenteraient aucune réponse cohérente aux questions pouvant être posées sur les causes de cette occupation.

48

Les recherches futures dans le nord de la Gaule pourraient s'attacher à réaliser une sélection drastique des sites participant de la conception et de l'évaluation des modèles, en association avec des reconstitutions paléoenvironnementales fondées sur de nouvelles études géomorphologiques. Enfin, divers facteurs sociaux, économiques et culturels pourraient encore être modélisés et mettre en évidence des tendances inattendues dans les choix d'implantation des établissements ruraux.

Bibliographie

Agache, Roger. 1978. *La Somme pré-romaine et romaine : d'après les prospections aériennes à basse altitude*. Amiens : Société des antiquaires de Picardie.

Bayard, Didier et Jean-Luc Collart. 1996. « De la ferme indigène à la villa romaine. La romanisation des campagnes de la Gaule ». *Revue archéologique de Picardie*. 11 : 5-8. <https://doi.org/10.3406/pica.1996.1883>.

Bayard, Didier, Nathalie Buchez et Pascal Depaepe, éd. 2014. *Quinze ans d'archéologie préventive sur les grands tracés linéaires en Picardie : Seconde partie*. Amiens : Université de Picardie.

Ben Redjeb, Tahar. 2013. *Carte archéologique de la Gaule : 80/2. La Somme*. Paris : Éditions de la Maison des sciences de l'homme.

Ben Redjeb, Tahar, Laurent Duvette, Pascal Quérel, Cyrille Chaidron, Sébastien Lepetz, Véronique Matteredne et Denis Maréchal. 2005. « Les campagnes antiques : bilans et perspectives ». *Revue archéologique de Picardie* 3-4 : 177-222. <https://doi.org/10.3406/pica.2005.2466>.

Blondiau, Lydie. 2014. « L'occupation rurale gallo-romaine de la vallée de La Selle : première approche » dans *Workshop 1 : Établissements ruraux du nord et de l'est de La Gaule : Données récentes*, édité par Michel Reddé, 1-15. Paris : École pratique des hautes études. <https://halshs.archives-ouvertes.fr/halshs-01058493>.

Deru, Xavier. 2012. « L'occupation du Cambrésis à la période romaine ». *Revue du Nord-Archéologie* 94 (398) : 121-141. <https://doi.org/10.3917/rdn.398.0121>.

Ferdière, Alain, Cristina Gandini, Pierre Nouvel et Jean-Luc Collart. 2010. « Les grandes villae "à pavillons multiples alignés" dans les provinces des Gaules et des Germanies : répartition, origine et fonctions ». *Revue archéologique de l'Est* 59 : 357-446. <http://journals.openedition.org/rae/6217>.

Gaffney, Vincent et Martijn van Leusen. 1995. « Postscript – GIS, Environmental Determinism and Archaeology : A Parallel Text » dans *Archaeology and Geographical Information Systems : A European Perspective*, édité par Gary R. Lock et Zoran Stančić, 367-382. London, Bristol : Taylor & Francis.

Gandini, Cristina, François Favory et Laure Nuninger, éd. 2012. *Settlement Pattern, Production and Trades from Neolithic to Middle Ages. ARCHAEDYN : 7 Millennia of Territorial Dynamics. Final Conference. University of Burgundy, Dijon, 23-25 June 2008*. Oxford : Archaeopress.

Güimil-Fariña, Alejandro et César Parcerro-Oubiña. 2015. « "Dotting the Joins" : A Non-Reconstructive Use of Least Cost Paths to Approach Ancient Roads. The Case of the Roman Roads in the NW Iberian Peninsula ». *Journal of Archaeological Science* 54 : 31-44. <https://doi.org/10.1016/j.jas.2014.11.030>.

Habermehl, Diederick. 2011. « Exploring Villa Development in the Northern Provinces of the Roman Empire » dans *Villa Landscape in the Roman North. Economy, Culture and Life-styles*, édité par Nico Roymans et Ton Derks, 64-83. Amsterdam : Amsterdam University Press.

Habermehl, Diederick. 2013. *Settling in a Changing World : Villa Development in the Northern Provinces of the Roman Empire*. Amsterdam : Amsterdam University Press.

Harris, Trevor et Gary Lock. 2006. « Enhancing Predictive Archaeological Modeling » dans *GIS and Archaeological Site Location Modeling*, édité par Mark W. Mehrer et Konnie L. Wescott, 41-62. Boca Raton : CRC Press. <https://doi.org/10.1201/9780203563359>.

Herzog, Irmela 2013. « Theory and Practice of Cost Functions » dans *Fusion of Cultures. Proceedings of the 38th Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Granada, Spain, April 2010*, édité par Francisco Contreras, Mercedes Farjas et Francisco Javier Melero, 375-382. Oxford : Archaeopress. https://publikationen.uni-tuebingen.de/xmlui/bitstream/handle/10900/60793/59_Herzog_CAA_2010.pdf?sequence=2&isAllowed=y.

Jochim, Michael A. 1976. *Hunter-Gatherer Subsistence and Settlement : A Predictive Model*. New York : Academic Press Inc.

Judge, W. James et Lynne Sebastian, éd. 1988. *Quantifying the Present and Predicting the Past : Theory, Method, and Application of Archaeological Predictive Modeling*. Denver, Washington : United States Bureau of Land Management. <http://archive.org/details/quantifyingprese00judg>.

Kamermans, Hans 2004. « The Application of Predictive Modelling in Archaeology : Problems and Possibilities » dans *Beyond the Artefact. Digital Interpretation of the Past. Proceedings of CAA2004. Prato 13-17 April 2004*, édité par Franco Nicolucci et Sorin Hermon, 273-277. Budapest : Archaeolingua. <https://openaccess.leidenuniv.nl/handle/1887/21011>.

Kohler, Timothy A. et Sandra C. Parker. 1986. « Predictive Models for Archaeological Resource Location ». *Advances in Archaeological Method and Theory* 9 : 397-452.

Kvamme, Kenneth L. 2006. « There and Back Again » dans *GIS and Archaeological Site Location Modeling*, édité par Mark W. Mehrer et Konnie L. Wescott, 3-38. Boca Raton : CRC Press. <https://doi.org/10.1201/9780203563359>.

Lepetz, Sébastien et Véronique Matterné. 2003. « Élevage et agriculture dans le nord de la Gaule durant l'époque gallo-romaine : une confrontation des données archéologiques et carpologiques ». *Revue archéologique de Picardie* 1 (1) : 23-35. <https://doi.org/10.3406/pica.2003.2354>.

Lock, Gary et John Pouncett. 2010. « Walking the Ridgeway Revisited : The Methodological and Theoretical Implications of Scale Dependency for the Derivation of Slope and the Calculation of Least-Cost Pathways » dans *Making History Interactive. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 37th International Conference, Williamsburg, Virginia, United States of America, March 22-26 2009*, édité par Bernard Frischer, Jane Webb Crawford et David Koller, 192-203. Oxford : Archaeopress. https://bibliographie.uni-tuebingen.de/xmlui/bitstream/handle/10900/61882/25_Lock_Pouncett_-CAA2009.pdf?sequence=2&isAllowed=y.

Matterné, Véronique. 2001. *Agriculture et alimentation végétale durant l'âge du fer et l'époque gallo-romaine en France septentrionale*. Autun : Éditions Mergoïl.

Nisard, Charles. 1864. *Les agronomes latins : Caton, Varron, Columelle, Palladius. Avec la traduction en français. Publiés sous la direction de M. Nisard*. Paris : Firmin-Didot Frères, Fils et Cie. <https://gallica.bnf.fr/ark:/12148/bpt6k2820632>.

Nuninger, Laure, Philip Verhagen, Frédérique Bertoncello et Angelo Castrorao Barba. 2016. « Estimating "Land Use Heritage" to Model Changes in Archaeological Settlement Patterns » dans *3rd International Landscape Archaeology Conference 2014 Rome, 17th – 20th of September 2014*, édité par Rita Hermans et Arjan de Rooy, 1-12. Amsterdam : VU E-Publishing. <https://doi.org/10.5463/lac.2014.60>.

Nüsslein, Antonin 2016. « Les campagnes entre Moselle et Rhin dans l'Antiquité. Analyse comparative des dynamiques spatiales et temporelles du peuplement de quatre micro-régions du 1^{er} siècle avant Jésus-Christ au 5^e siècle après Jésus-Christ. » Thèse de doctorat, Strasbourg : Université de Strasbourg. <http://www.theses.fr/s78127>.

Stančić, Zoran et Kenneth L. Kvamme. 1999. « Settlement Pattern Modelling through Boolean Overlays of Social and Environmental Variables » dans *New Techniques for Old Times. CAA98. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 26th Conference, Barcelona, March 1998*, édité par Juan A. Barceló, Ivan Briz et Asumpcio Vila, 757 : 231-238. Oxford : Archaeopress. http://proceedings.caaconference.org/files/1998/36_Stancic_Kvamme_CAA_1998.pdf.

Tobler, Waldo. 1993. « Three Presentations on Geographical Analysis and Modeling : Non-Isotropic Geographic Modeling ; Speculations on the Geometry of Geography ; and Global Spatial Analysis ». Technical Report 1. Santa Barbara : National Center for Geographic Information and Analysis. <https://escholarship.org/uc/item/05r820mz>.

Verhagen, Philip et Karen Jeneson. 2012. « A Roman Puzzle. Trying to Find the via Belgica with GIS » dans *Thinking Beyond the Tool. Archaeological Computing and the Interpretive Process. Proceedings of the 32nd Annual Meeting of the Theoretical Archaeology Group, 17-19 December 2010, Bristol*, édité par Angeliki Chrysanthi, Patricia Murrieta-Flores et Constantinos Papadopoulos, 123-130. Oxford : Archaeopress. <https://research.vu.nl/files/2947935/Verhagen%20Jeneson%202012.pdf>.

Verhagen, Philip, Laure Nuninger, Frédérique Bertoncello, François-Pierre Tourneux et Karen Jeneson. 2013. « Introducing the Human Factor in Predictive Modelling : A Work in Progress » dans *Archaeology in the Digital Era. Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton, 26-29 March 2012*, édité par Earl Graeme, Tim Sly, Angeliki Chrysanthi, Patricia Murrieta-Flores, Constantinos Papadopoulos, Iza Romanowska et David W. Wheatley, 379-388. Amsterdam : Amsterdam University Press. <https://halshs.archives-ouvertes.fr/halshs-00743700>.

Verhagen, Philip, Laure Nuninger, Frédérique Bertoncello et Angelo Castrorao Barba. 2015. « Estimating the “Memory of Landscape” to Predict Changes in Archaeological Settlement Patterns » dans *Keep the Revolution Going. 43rd Computer Applications and Quantitative Methods in Archaeology Conference (CAA 2015 SIENA), Mar 2015, Siena, Italy*, édité par Stefano Campana, Roberto Scopigno, Gabriella Carpentiero et Marianna Cirillo, 623-636. Oxford : Archaeopress. <https://hal.archives-ouvertes.fr/hal-01252715/document>.

Verhagen, Philip et Thomas G. Whitley. 2012. « Integrating Archaeological Theory and Predictive Modeling : A Live Report from the Scene ». *Journal of Archaeological Method and Theory* 19 (1) : 49-100. <https://doi.org/10.1007/s10816-011-9102-7>.

Wheatley, David. 2004. « Making Space for an Archaeology of Place ». *Internet Archaeology* 15 : 1-22. <https://doi.org/10.11141/ia.15.10>.

Whitley, Thomas G. 2004a. « Causality and Cross-Purposes in Archaeological Predictive Modeling » dans *[Enter the Past.] The E-way into the Four Dimensions of Cultural Heritage. CAA 2003 : Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 31st Conference, Vienna, Austria, April 2003*, édité par Karin Fischer-Ausserer, Wolfgang Börner, Maximilian Goriany et Lisa Karlhuber-Vöckl, 1227 : 1-17. Oxford : Archaeopress. https://publikationen.uni-tuebingen.de/xmlui/bitstream/handle/10900/62123/CD49_Whitley_CAA_2003.pdf?sequence=2.

Whitley, Thomas G. 2004b. « Spatial Variables as Proxies for Modeling Cognition and Decision-Making in Archaeological Settings : A Theoretical Perspective ». *Internet Archaeology* 16 : 1-16. <https://doi.org/10.11141/ia.16.3>.

Wightman, Edith Mary. 1985. *Gallia Belgica*. Berkeley : University of California Press.

Notes

1 Traitement logiciel réalisé avec ArcMap 10.4 (© ESRI).

Auteur

Nicolas Revert

Pôle Archéologie préventive, Metz Métropole, France

Archéologue au sein de Metz Métropole où il participe à la conception et à l'enrichissement d'un SIG de service, Nicolas Revert a réalisé des recherches en modélisation spatiale et prédictive par le biais d'un master d'archéologie numérique à l'université de Leyde aux Pays-Bas, concernant l'occupation romaine de la Somme.

ORCID 0000-0002-5169-604X

nicolas.revert.arch@gmail.com

Droits d'auteur

Les contenus de la revue *Humanités numériques* sont mis à disposition selon les termes de la [Licence Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/).