
Exploitation des données de diagnostics en tranchées mécaniques par l'analyse spatiale

Spatial analysis for an efficient use of data from trial trenching

Sylvain Badey et Xavier Rodier



Édition électronique

URL : <https://journals.openedition.org/racf/2202>

ISSN : 1951-6207

Éditeur

Fédération pour l'édition de la Revue archéologique du centre de la France (FERACF)

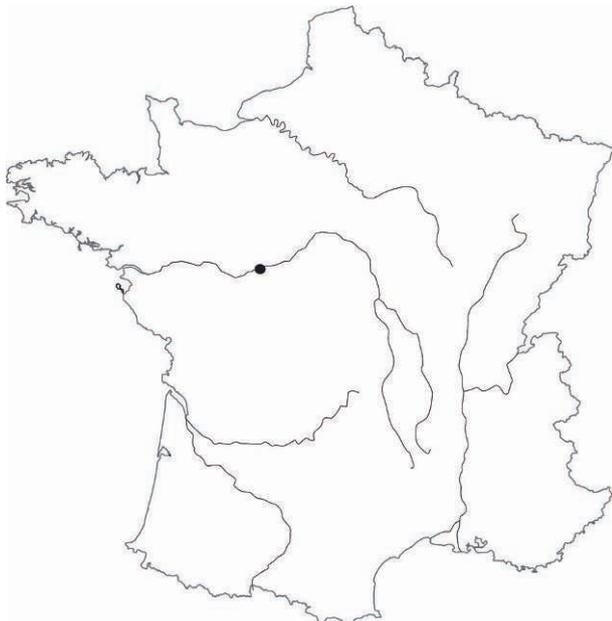
Référence électronique

Sylvain Badey et Xavier Rodier, « Exploitation des données de diagnostics en tranchées mécaniques par l'analyse spatiale », *Revue archéologique du Centre de la France* [En ligne], Tome 53 | 2014, mis en ligne le 30 juin 2015, consulté le 06 avril 2023. URL : <http://journals.openedition.org/racf/2202>



Creative Commons - Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International
- CC BY-NC-ND 4.0

<https://creativecommons.org/licenses/by-nc-nd/4.0/>



Sylvain BADEY* et Xavier RODIER**

Exploitation des données de diagnostics en tranchées mécaniques par l'analyse spatiale

SPATIAL ANALYSIS FOR AN EFFICIENT USE OF DATA FROM TRIAL TRENCHING

Mots-clés : Diagnostic archéologique, analyse spatiale, tranchées, occupation du sol, prescription archéologique.

Keywords: *Preventive archeology, archaeological assessment, spatial analysis, trial trenches, land use, archaeological prescription.*

Résumé : Les diagnostics archéologiques en tranchées mécaniques sont effectués pour fournir au service de l'État les éléments tangibles d'aide à la décision en vue de la prescription. Les surfaces concernées par ces diagnostics sont considérables et livrent des informations diachroniques sur l'occupation du sol dont l'exploitation est souvent limitée au cadre des fouilles prescrites par la suite. L'article présente les méthodes et les résultats d'une expérience menée à partir des données du diagnostic de la ZAC d'Ozans à Étretchet dans l'Indre. Il s'agissait d'évaluer l'apport de l'analyse spatiale, d'une part à la lecture des résultats du diagnostic, d'autre part à la compréhension de l'occupation sol dans le secteur couvert. L'utilisation de méthodes d'analyse spatiale exploratoire montre que, sans remplacer l'analyse archéologique indispensable, il est possible de tester des hypothèses interprétatives sur la répartition des vestiges en s'affranchissant du découpage imposé par les tranchées. Les résultats obtenus sont utiles à l'interprétation des vestiges et au-delà constituent une aide potentielle à la prescription.

*. Inrap – UMR 7324 CITERES LAT, Université François Rabelais – CNRS, Tours.

** . Inrap – UMR 7324 CITERES LAT, Université François Rabelais – CNRS, Tours.

Pour citer cet article, utiliser la référence électronique :

Sylvain Badey et Xavier Rodier, Exploitation des données de diagnostics en tranchées mécaniques par l'analyse spatiale, *Revue archéologique du Centre de la France* [En ligne], Tome 53 | 2014, mis en ligne le 16 avril 2015, consulté le 19 juin 2015.
URL : <http://racf.revues.org/2202>

Abstract: *The archaeological diagnostic consists in assessing the archaeological potential of an area impacted by development. This first phase aims to provide sufficient information -on the nature, extent, chronology and degree of preservation of the sites- to the competent authorities (in France, the Ministry of Culture) in view to prescribe or not an excavation. Beyond the assessment of archaeological potential, they deliver diachronic information on land-use whose exploitation is often restricted. In the case of large areas surveyed, we let's suppose that the data mining is useful for archaeological interpretation of the site. This paper presents the methods and results of an experiment based on data from a 226 ha archaeological survey performed at Etrechet in Indre (Centre of France). The aim was to evaluate the contribution of spatial analysis, on one hand to read the results of the trial trenching, secondly to understand the land-use in the studied area. The experiment focused on the use of exploratory spatial data analysis. The methods used are both cartographic with the grid analysis and statistic with data clustering. The aim of this methods is to overcome the division imposed by the trial trenches, and consider the whole area. Results show that it is possible to test interpretive hypothesis on the distribution of the remains. Indeed spatial analysis reveals patterns by periods, confronted to the essential archaeological interpretation. The results meet the two objectives of an archaeological assessment. First, they are useful to interpret the remains and to understand the dynamics of land-use. Secondly, as decision-making tool, they provide elements for the authorities to deal with the archaeological sites (excavations or preservation).*

INTRODUCTION

1. ANALYSE SPATIALE

DE LA CONCENTRATION DES VESTIGES

1.1. OUTILS ET PROCESSUS : LA MOBILISATION DES DONNÉES

1.2. OUTILS ET PROCESSUS : LES NIVEAUX D'ANALYSE

1.3. LES DONNÉES PRISES EN COMPTE : LE CORPUS

1.4. ANALYSE PAR MAILLE

1.5. MODÉLISATION ET INDICES DE FAITS

1.6. LES RÉSULTATS DE L'ANALYSE PAR MAILLE

2. ANALYSE SPATIALE EXPLORATOIRE

2.1. LA MÉTHODE DU VOISIN LE PLUS PROCHE

2.2. LES RÉSULTATS

2.3. LES FONCTIONS K DE RIPLEY ET L DE BESAG

2.4. LES RÉSULTATS

3. REPRÉSENTATION DES AGRÉGATS AVEC UN ALGORITHME DE SCAN SPATIAL : LA MÉTHODE DES FENÊTRES MOBILES

LES RÉSULTATS

CONCLUSION

■ BIBLIOGRAPHIE

INTRODUCTION

Le diagnostic archéologique en tranchées mécaniques est la méthode prospective la plus communément employée pour fournir à l'État les éléments tangibles d'aide à la décision pour l'établissement de prescriptions de fouilles archéologiques préventives. Elle permet de mettre en évidence les vestiges archéologiques, de les caractériser et de les restituer dans un contexte historique et environnemental régional.

Selon le Livre blanc de l'archéologie rédigé en 2013, sur les onze dernières années (2002-2012), on estime à 60 000 ha les surfaces aménagées en France chaque année. 20 % de celles-ci font l'objet de diagnostics archéologiques soit 2 200 par an, en moyenne. Ses opérations donnent lieu en moyenne à la réalisation de 480 fouilles.

Les diagnostics en tranchées mécaniques sont à l'origine de la majorité des données archéologiques produites chaque année et servent à la fois à la prescription de fouilles et à l'alimentation de la carte archéologique. L'objet de cette étude¹ comporte deux champs d'exploration espérant servir d'une part à améliorer la lecture des résultats issus d'un diagnostic archéologique sur une grande surface pour les prescripteurs, d'autre part à enrichir la compréhension de l'occupation du sol grâce aux données archéologiques produites, bien qu'elles soient parcellaires.

L'utilisation de méthodes issues de l'analyse spatiale exploratoire n'a pas vocation à remplacer l'analyse archéologique mais au contraire à la consolider par la confrontation des interprétations issues du terrain et des résultats des traitements de données.

Les diagnostics effectués sous la responsabilité de Matthieu Munos et Nicolas Holzem (Inrap) entre janvier 2009 et janvier 2011 en amont du projet d'aménagement de la ZAC d'Ozans sur la commune d'Étrechet (Indre) couvrent une surface totale de 226 ha. (Fig. 1) Cette première phase du diagnostic de la ZAC d'Ozans a été prescrite en cinq tranches qui ont été diagnostiquées en tranchées mécaniques continues séparées de 18 mètres bord à bord. En outre, des fenêtres d'extensions ont été ouvertes quand cela était jugé nécessaire pour la compréhension des vestiges mis au jour. Le taux d'ouverture total des tranchées représente 11 % de la surface

prescrite. La mise en place d'un système d'information géographique (SIG) par Denis Godignon (Inrap), dès cette phase 1 du diagnostic a permis de disposer de données structurées pour les analyses spatiales de cette étude. Elle visait à rendre plus pertinente la lecture des données du diagnostic pour l'établissement des prescriptions de fouille (aide à la décision) ainsi qu'à améliorer la connaissance de l'occupation du sol sur la zone diagnostiquée. Afin de ne pas biaiser l'analyse, seules les données du diagnostic acquises sur le terrain et étudiées en post-fouille ont été utilisées sans tenir compte des résultats des fouilles qui ont suivi.

La pertinence des méthodes de diagnostic en tranchées et les facteurs influant sur la capacité à détecter des vestiges archéologiques par ce moyen ont déjà été plusieurs fois étudiés et discutés (BLANCQUAERT et MEDLYCOTT 2006 ; BRUN et VANMOERKERKE 2006 ; DUBOULOZ 2003). La stratégie d'implantation des tranchées mécaniques et leur efficacité a notamment fait l'objet d'un article de P. Verhagen et A. Borsboom (2009) dans lequel ils confrontent les différentes configurations de tranchées grâce à la notion de probabilité d'intersection.

1. ANALYSE SPATIALE DE LA CONCENTRATION DES VESTIGES

L'analyse spatiale a pour objectif " de décrire une disposition particulière de certains objets, leur organisation spatiale, repérer des structures, expliquer une localisation par d'autres [...] déceler en quoi la localisation apporte un élément utile à la connaissance des objets étudiés et peut en expliquer les caractéristiques, en totalité ou en partie " (PUMAIN et SAINT-JULIEN 2010). Il faut donc préalablement décrire les propriétés spatiales des objets en question et préciser les échelles d'analyses qui seront utilisées.

1.1. Outils et processus : la mobilisation des données

La première étape consiste à regrouper les données archéologiques acquises sur le terrain puis lors de l'étude post-fouille pour pouvoir disposer d'un corpus homogène. Cette tâche a été simplifiée par la formalisation, dans une base de données et un SIG, dès la phase de terrain à l'aide d'une tablette PC, d'un enregistrement normalisé des données du diagnostic.

Les données archéologiques utilisées sont réunies dans une base de données spatiale : les polygones représentant l'emprise de diagnostic, celle

1. Cette étude de cas a été réalisée dans le cadre d'un rapport de stage au Laboratoire Archéologie et Territoires de l'UMR 7324 CITERES au cours d'une année de Master 2 pro à l'Université François Rabelais de Tours et soutenu en juin 2012.



Fig. 1 : Vue générale de l'emprise de la zone d'étude, des tranches de diagnostic prescrites et des tranchées ; fond Bing maps, été 2011 avec les fouilles réalisées.

des ouvertures et enfin les Faits archéologiques. Les données descriptives issues de l'enregistrement archéologique de terrain et des études postérieures sont associées à chaque entité. En effet, pour les analyses spatiales ce sont les propriétés descriptives (identification, géométrie), géographiques (localisation) et chronologiques (datation) des vestiges archéologiques qui caractérisent l'objet d'étude.

1.2. Outils et processus : les niveaux d'analyse

L'étude archéologique se fait, par essence, à des niveaux d'analyses et d'interprétation différents et imbriqués. Les trois niveaux que nous avons jugés pertinents sont définis en fonction de l'échelle d'analyse et du processus d'acquisition et d'interprétation des données archéologiques (Fig. 2). Ils sont

évidemment transposables à d'autres opérations de diagnostic archéologique.

Le premier, à la plus petite échelle, est imposé par l'étendue de la zone d'étude et des données acquises sur le terrain : c'est celui de l'emprise globale du diagnostic. Les vestiges peuvent y être étudiés dans leur spécificité géographique, grâce à l'utilisation d'un maillage qui traduit la répartition des Faits archéologiques par exemple.

Le deuxième, qui correspond à la même échelle, distingue les vestiges selon leur attribution chronologique (périodes). Il permet d'étudier la répartition et les concentrations de Faits d'une part avec l'utilisation d'un maillage, d'autre part via l'analyse exploratoire. C'est sur ce niveau que porte l'étude.

Les résultats des analyses et leur confrontation avec les interprétations archéologiques permettent de définir des zones de concentrations de vestiges ou

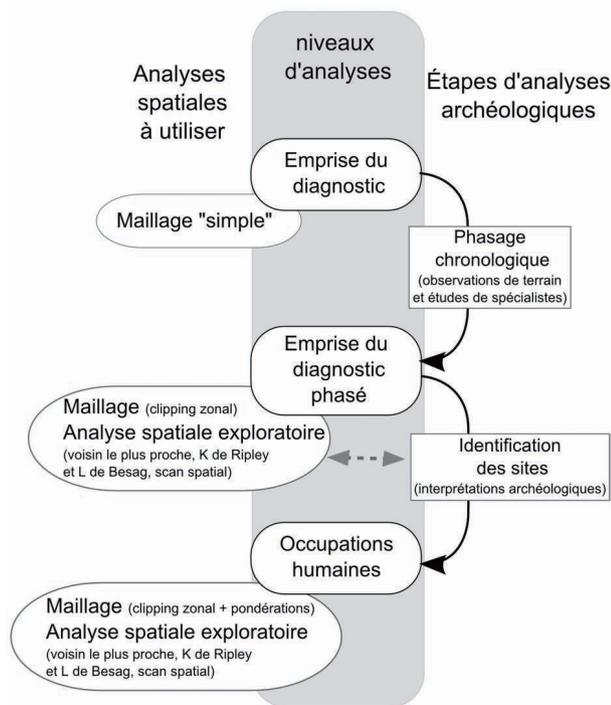


Fig. 2 : Niveaux d'analyse et outils mis en œuvre pour l'étude de l'espace diagnostiqué.

des occupations humaines par périodes et amènent naturellement à la définition d'un troisième niveau d'analyse, celui des occupations humaines. Des méthodes similaires ou proches de celles mises en œuvre ici pourraient être appliquées à cette plus grande échelle.

1.3. Les données prises en compte : le corpus

Le choix de l'élément de base pour l'analyse spatiale est le Fait archéologique dans une acception très large du terme qui correspond à l'unité d'enregistrement du terrain et qui peut être synonyme de vestige archéologique. Leur enregistrement dans un tableau de Faits est accompagné de nombreux attributs qualitatifs et quantitatifs de natures variées : administratives (numéro de parcelle, de phase...), géométriques (diamètre/longueur, largeur et profondeur) et archéologiques (interprétation, type de mobilier, datation...).

Plusieurs critères sont utilisés pour la datation des vestiges archéologiques mis au jour lors d'un

diagnostic : typo-chronologie du mobilier qu'il recèle, relations stratigraphiques ou topologiques entre Faits, sources iconographiques anciennes. Il en résulte une multiplicité des qualités de précisions chronologiques qui devait être simplifiées. Les datations ont donc fait l'objet d'une reclassification dans un champ appelé PERIODES afin de les regrouper par grandes classes (Fig. 3). Quatre "périodes" ont été retenues pour l'analyse : la Protohistoire, l'Antiquité, le Moyen Âge et l'époque Moderne.

1.4. Analyse par maille

Pour étudier la répartition par périodes des vestiges archéologiques sur la totalité de la zone et émettre des hypothèses quant aux disparités de l'occupation perceptibles chronologiquement et spatialement, nous avons dans un premier temps réalisé une analyse par maille à l'aide d'un carroyage.

L'utilisation d'une analyse par maille (ou tessellation) permet de s'affranchir, dans une certaine mesure, des limites de l'implantation des tranchées qui, si elle est relativement régulière crée néanmoins un masque pour la représentation des vestiges archéologiques. L'utilisation d'une maille carrée de 50 m de côté, dans laquelle on dénombre les Faits, permet de s'affranchir de cette dernière en intersectant au minimum deux tranchées et correspond au niveau d'analyse fixé.

Le choix du carroyage définit la résolution d'étude et de représentation de la zone de recherche qui est ainsi découpée en 1 024 carrés de 50 m de côté dont 87 seulement ne sont pas intersectés par une tranchée, majoritairement en bord d'emprise.

1.5. Modélisation et indices de Faits

Afin de représenter l'importance des vestiges d'une même période compris dans chaque maille, le comptage simple ne tient pas compte des Faits archéologiques présents dans plusieurs mailles. Il convient alors de passer d'un carroyage adressé à un carroyage ventilé tels que le définit G. Lajoie (LAJOIE et BUSSI 1993). Pour la répartition des données nous avons utilisé l'algorithme de *clipping* de zone qui les ventile selon un principe purement géométrique : le pourcentage de chaque Fait compté par cellule est fonction de la surface du Fait porté par chaque cellule du carroyage (LAJOIE 1991 ; LAJOIE et BUSSI 1993). Enfin, la somme des proportions surfaciques de Faits inclus dans une maille est pon-

PERIODES	valeurs saisies au départ	Nombre de "faits"	dont Mobilier Isolé	Échantillon retenu
PREHISTOIRE	Prehistoire, Néolithique	17	14	3
PROTOHISTOIRE	Neo_Proto, Age du Bronze, Bronze Final, Age du Fer, La Tene Ancienne, La Tene, La Tene Finale, Hallstat, Protohistoire	279	74	205
ANTIQUE	La Tene Finale_GR, Ier-IIIe s., Antiquité tardive, GR	174	28	146
MOYEN AGE	Haut MA, MA, MA Classique, Debut MA Classique, Bas MA, Bas MA_Debut MOD, MA ou MO	99	25	74
EPOQUE MODERNE	Debut MOD, Moderne, Moderne ou Contemporain	138	8	130
NON DATE	non daté	1513	148	1365
TOTAL		2220	297	1923

Fig. 3 : Reclassification des valeurs saisies au départ dans le champ " PERIODES " et nombre de Faits pris en compte dans l'échantillon statistique (hors Faits annulés ou contemporains).

dérée par le nombre total de Faits par périodes afin d'obtenir une répartition de classes plus homogène entre les différentes périodes (Fig. 4). Cette pondération est d'autant plus nécessaire que les périodes présentent de fortes disparités en nombre de Faits, dues à la fois aux propriétés intrinsèques des vestiges – une *villa* gallo-romaine laissant des traces plus nombreuses et plus grandes qu'un petit habitat sur poteaux – mais aussi à l'importance de la représentation de chaque période – les vestiges médiévaux étant moins nombreux que ceux de l'Antiquité et de la Protohistoire.

Le taux, ou indice, obtenu par maille, a fait l'objet d'une classification à l'aide de la méthode des seuils naturels (*Jenks*) que nous avons adaptée afin de rendre les résultats comparables d'une période à l'autre. Six classes ont été définies de 0,1 % à 0,5 % de surface de Faits par maille par rapport à la totalité (gris clair), jusqu'à l'intervalle de 5 % à 35 % considéré comme étant une proportion importante de Faits d'une période dans une maille (noir). La borne supérieure de la classe supérieure est différente selon la période étudiée. C'est pourquoi la Fig. 5 indique un intervalle générique à toutes les cartes alors que les Fig. 6 à 9 montrent les bornes strictes de chaque période associées aux statistiques descriptives simples de l'échantillon concerné.

Enfin, pour représenter les changements de répartition des vestiges entre deux périodes, la valeur de chaque maille pour la période la plus ancienne est retranchée à celle des mailles correspondantes de la période la plus récente (Fig. 5). Les valeurs sont ici aussi divisées en six classes à partir de la méthode

des seuils naturels adaptée pour rendre les résultats comparables d'une carte à l'autre : de l'intervalle -30 à -3 (bleu foncé), qui représente une dépréciation de l'indice, c'est-à-dire une baisse de la proportion de Faits de la période la plus récente par rapport à la période la plus ancienne, jusqu'à l'intervalle 3 à 35 (rouge foncé) indiquant, à l'inverse, une augmentation de l'indice.

Comme pour l'analyse par maille, les bornes inférieures de la plus petite classe et les bornes supérieures de la plus grande varient selon la période. La légende de la Fig. 5 indique les classes génériques alors que les extractions de cartes par périodes deux à deux (Fig. 10 à 12) montrent les valeurs strictes de l'échantillon ainsi que les statistiques simples de chaque échantillon.

1.6. Les résultats de l'analyse par maille

Les résultats de la Phase 1 du diagnostic de la ZAC d'Ozans que nous comparons ici à l'analyse par maille sont tous issus des cinq rapports finaux d'opérations (MUNOS 2010a, b, c et d ; HOLZEM 2011) .

1.6.1. Analyse par maille de la proportion de Faits par périodes

– La Protohistoire (Fig. 6)

Cette période est la plus représentée dans la zone d'étude avec 205 Faits comptabilisés pour cette ana-

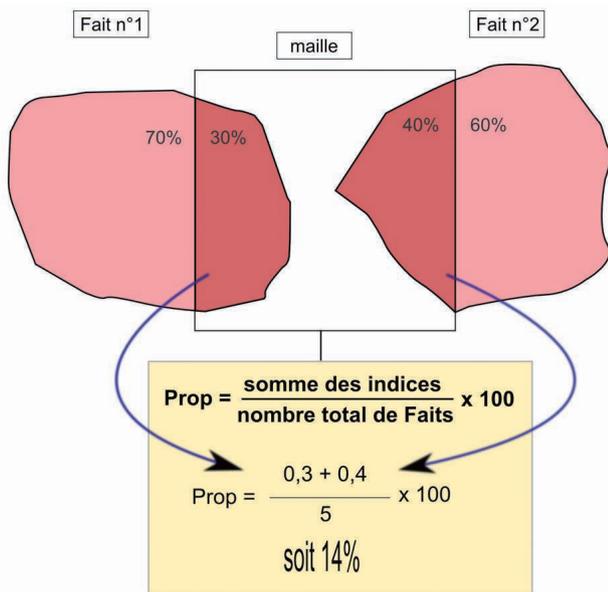


Fig. 4 : Exemple de calcul de la proportion de Faits comptée dans une maille pour un total théorique de 5 Faits.

lyse, ceci tient en partie au fait qu'elle s'étend sur près de deux millénaires, de l'âge du Bronze à la fin de l'âge du Fer. Des vestiges de cette période sont présents dans les quatre tranches du diagnostic, et malgré une dispersion apparente se regroupent en quatre zones remarquables sur la représentation par maille. En effet, si l'on s'attache aux mailles ayant les taux les plus élevés, de 1 à 7 %, et sur les ensembles qu'ils forment, on peut distinguer du sud vers le nord :

- Au sud-ouest, dans la tranche 1, un ensemble de mailles dont les taux sont moyen à fort, 0,5 à 3 % est entouré de carreaux à taux plus faible, 0,1 à 0,5 %. Cet ensemble assez disparate laisse néanmoins transparaître une présence importante de Faits dans cette zone, qui pourrait correspondre à la présence d'une ou plusieurs occupations protohistoriques datées de l'âge du Bronze final, d'après le mobilier archéologique. Ces concentrations de Faits relativement importantes, mais sans organisation apparente, sont la trace d'une ou plusieurs occupations protohistoriques mais dont l'organisation reste difficilement compréhensible.
- Un peu plus au nord, dans la tranche 5, un ensemble de mailles très concentrées présente des taux de Faits protohistoriques de 0,1 à 7 %. Il s'agit de bâtiments laténiens sur poteaux,

représentés par de plus fortes concentrations. D'autres trous de poteau, des grandes fosses d'extraction et des concentrations de mobilier de la même époque ou de l'âge du Bronze final, complètent cet ensemble et renforce l'hypothèse d'une occupation protohistorique.

- Au centre de la zone d'étude, à la frontière entre les tranches 3 et 4, plusieurs mailles correspondent à la découverte de nombreux vestiges protohistoriques. Si cet ensemble de mailles présente majoritairement des concentrations faibles, 0,1 à 1 %, il contient ponctuellement des cellules de plus fortes valeurs, 1 à 7 %, assez régulièrement répartis. Les mailles à fortes valeurs, 5 à 7 % des Faits protohistoriques du diagnostic, sont dues, respectivement, au nord-est à la présence d'un bâtiment sur poteaux de l'époque laténienne, au sud-est à la concentration de trous de poteau et de fosses ayant servi de structures de rejets secondaires et dont la céramique est datée de la fin du premier âge du Fer. La présence de bâtiments sur poteaux se traduit par des valeurs d'indices comprises entre 3 et 5 %. Les autres mailles à valeurs positives sont dues à la présence de fosses, de trous de poteau isolés ou de fossés d'enclos.
- Enfin au nord de la zone d'étude, dans la tranche 2 du diagnostic, les Faits protohistoriques contenus dans les mailles sont relativement dispersés. Néanmoins une densité de Faits plus marquée est visible au nord de l'ensemble. Ces Faits, mêlés à d'autres d'époques antérieures et postérieures, ne permettent pas de définir une organisation de l'occupation protohistorique dans cette zone.

L'analyse par maille des vestiges protohistoriques semble pertinente. En effet, les occupations découvertes sur le terrain sont toutes représentées avec leurs particularités : les bâtiments sur trous de poteau, composés d'au moins trois ou quatre Faits, apparaissent dans des mailles avec un fort taux, parfois proches de mailles à taux plus faible (fosses, fossés ou trous de poteau dispersés), formant des ensembles identifiables. De même, les zones d'occupations protohistoriques, dont l'organisation n'a pu être comprise, sont représentées par des ensembles de mailles à valeurs majoritairement positives mais rarement très fortes (5 à 7 % des Faits protohistoriques de l'échantillon).

L'analyse des vestiges de cette période est donc à faire à l'échelle de plusieurs mailles et non à celle d'une seule d'entre elle. Le choix de la maille, ici un carré de 50 m de côté soit une surface de 2 500 m²,

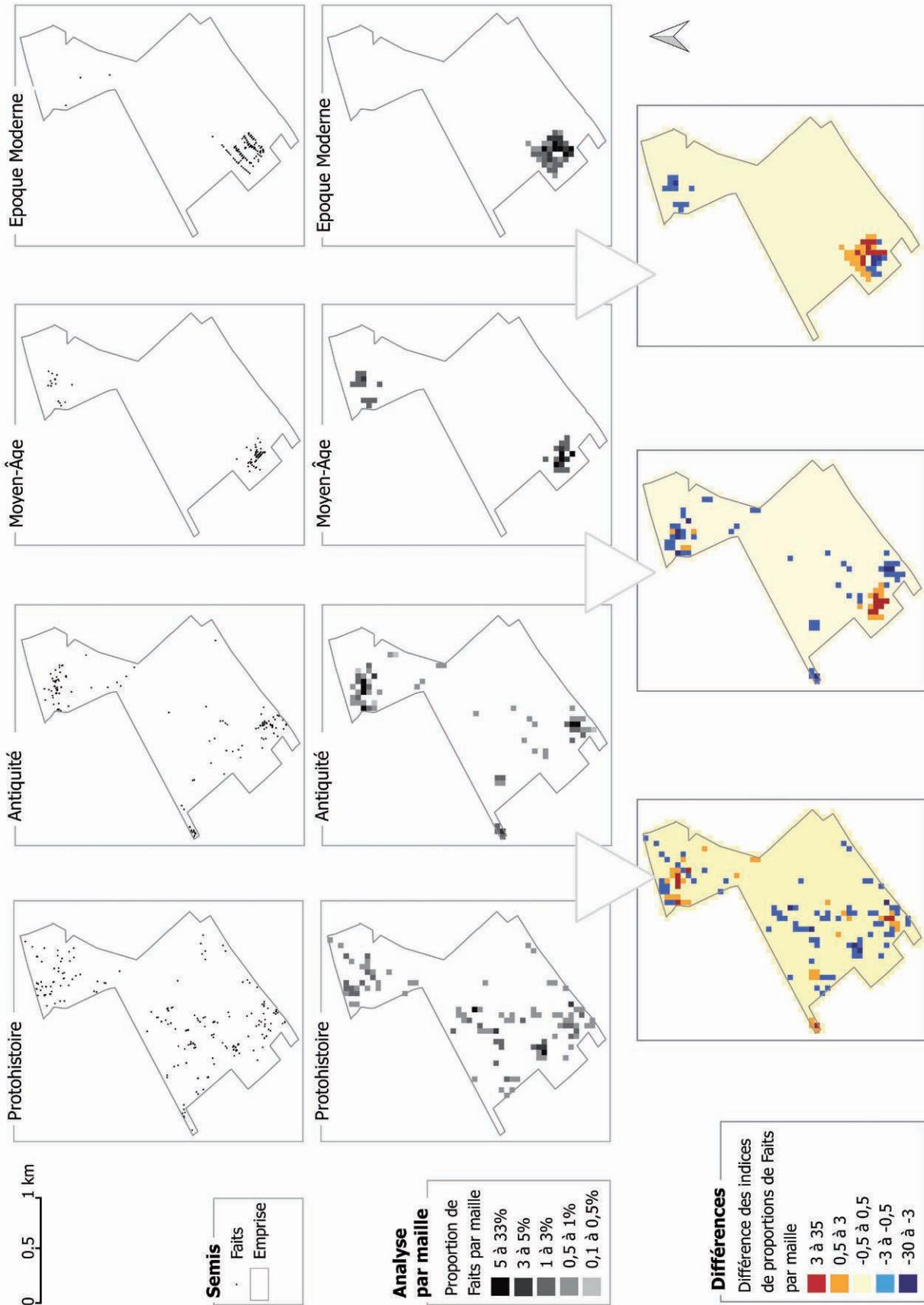


Fig. 5 : Semis de points, analyse par maille de l'occupation par périodes et cartes de différences inter-périodes.

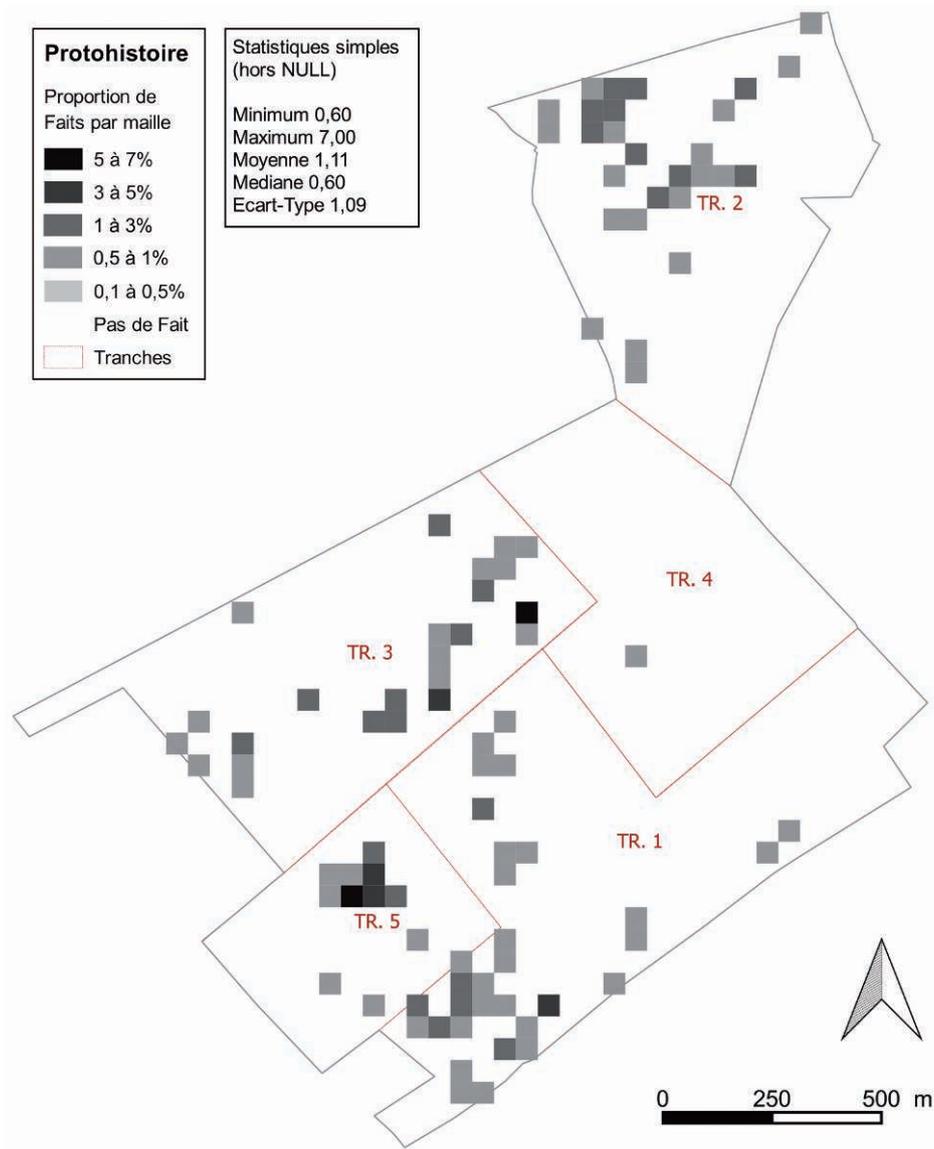


Fig. 6 : Analyse par maille de la proportion de Faits de la Protohistoire.

défini la taille minimale des sites reconnaissables et donne une lecture de leur organisation générale. On notera que le calcul de la proportion de Faits dans chaque maille ne tient compte que d'une partie des vestiges de la période. En effet, seuls les Faits datés sont pris en compte, c'est-à-dire essentiellement ceux qui contiennent du mobilier archéologique. Ainsi, un bâtiment composé de quinze trous de poteau mais dont un seul Fait a livré une quantité importante de céramiques (139 tessons) n'apparaît qu'au travers d'une maille à faible valeur.

En outre, les Faits sont comptabilisés dans leur "unicité". Par conséquent, chaque Fait n'est considéré que comme un seul élément quelle que soit sa taille. Les fossés et grandes fosses sont donc parfois dépréciés puisqu'un fossé qui s'étend sur quatre mailles sera pris en compte comme un quart de Fait, par exemple, dans chacune d'elles.

En définitive, l'occupation protohistorique est attestée par le diagnostic pour différentes périodes, de l'âge du Bronze à La Tène finale. Les zones identifiées par l'analyse par maille sont celles qui ont fait l'objet d'une prescription de fouille.

– L'Antiquité (Fig. 7)

L'Antiquité est également bien représentée avec 146 Faits datés, hors mobilier isolé, utilisés pour l'analyse par maille. Si c'est la période la plus courte telle que nous l'avons définie, elle est aussi sûrement la plus facile à caractériser. En effet, le matériel est bien identifiable (céramique et terres cuites architecturales par exemple) et les structures (*villae* maçonnées, fours à chaux ou de tuiliers, etc.) le sont souvent tout autant. Deux grandes zones sont perceptibles et remarquables à la fois par leur forte proportion de Faits par maille et par l'assemblage de cellules contiguës ou proches, l'une au sud-ouest dans la tranche 1 avec 15 mailles représentant de 0,5 à 14,7 % du nombre total de Faits, l'autre au nord, dans la tranche 2 avec 24 mailles représentant les mêmes valeurs de proportion de Faits.

- Au sud-ouest de la zone d'étude, une concentration de mailles centrée sur deux cellules à forte valeurs, entre 5 et 14,7 % de la totalité des Faits gallo-romains trouvés sur le diagnostic, est entourée de mailles de densité plus faible. Il se développe sur cet espace une *villa* maçonnée d'une surface estimée à 1 500 m² avec une ou deux constructions annexes. Celle-ci est entourée d'un atelier de potier (habitat, four et atelier), d'une carrière d'extraction de calcaire, d'un four à chaux et d'un probable four de tuilier. Cette occupation s'étend de la fin du I^{er} s. av. J.-C. jusqu'au III^e s. ap. J.-C. et occupe une surface d'environ trois hectares, soit douze mailles. La concentration importante et le noyau central de l'occupation qui se dégage du carroyage correspondent à ceux perçus lors du diagnostic.
- Au nord du diagnostic, dans la tranche 2, une autre concentration plus étendue apparaît. Elle est aussi centrée sur une ligne de trois cellules à forte proportion de Faits avec un ensemble de mailles présentant des taux moindres tout autour. Cet agrégat est le reflet d'une forte concentration de Faits antiques, notamment des ensembles maçonnés ou sur poteaux, ainsi que des structures artisanales – four et traitement de minerai de fer – qui sont compris dans les cellules aux taux les plus élevés. D'autres vestiges, plus dispersés et dont l'organisation spatiale est plus difficilement interprétable, complètent cette zone de concentration. Enfin, les mailles les plus à l'ouest traduisent probablement la densité de Faits d'un autre ensemble de la même période composé de murets d'en-

clos. Ce site pourrait se poursuivre à l'ouest au-delà de l'emprise du diagnostic.

- Enfin, un autre groupe de mailles présentant une forte proportion de Faits antique est située à l'ouest, à l'extrémité de la tranche 3. Sa situation enclavée dans l'emprise du diagnostic limite évidemment les interprétations spatiales en l'absence d'investigation aux alentours. Il s'agit de plusieurs Faits ayant livré des collections abondantes de mobilier (céramique, monnaies, objets métalliques, etc.) dans une fourchette chronologique allant du I^{er} au IV^e s. ap. J.-C.

L'analyse des proportions de Faits antiques par le carroyage est ici concordante avec les interprétations et les hypothèses émises par les archéologues. C'est la structuration des sites de cette période qui facilite cette adéquation et la lisibilité de l'organisation globale de ces occupations. C'est aussi probablement l'importance du matériel dans les Faits qui permet de les dater individuellement et d'obtenir une représentation moins lacunaire que pour d'autres périodes.

– Le Moyen Âge (Fig. 8)

Cette période est moins représentée que les précédentes avec 74 Faits datés, hors mobilier isolé, qui sont répartis en deux zones distinctes et relativement homogènes au sud-ouest et au nord de l'emprise du diagnostic.

- Dans la tranche 5, le premier ensemble s'étend sur une grande partie de la zone étudiée et comprend des mailles, presque toutes contiguës, aux proportions de Faits importantes. Cette concentration de Faits médiévaux traduit une forte occupation du sol étroitement liée à la seigneurie d'Ozans, dont le château actuel se situe un peu plus au sud-ouest (visible sur la photo aérienne utilisée en fond de carte, dans l'angle sud-ouest de la tranche 5 – Fig. 1). Les concentrations les plus fortes, de 5 à 34,9 % des Faits médiévaux découverts sur la totalité du diagnostic, sont centrées sur un fossé d'enceinte délimitant probablement la basse-cour du château et entourant une couche anthropisée contenant des concentrations importantes de matériel archéologique et des vestiges de maçonnerie. Les mailles les entourant avec des taux moindres mais néanmoins élevés, 1 à 3 % de la totalité des Faits de la période, traduisent la présence de Faits moins nombreux

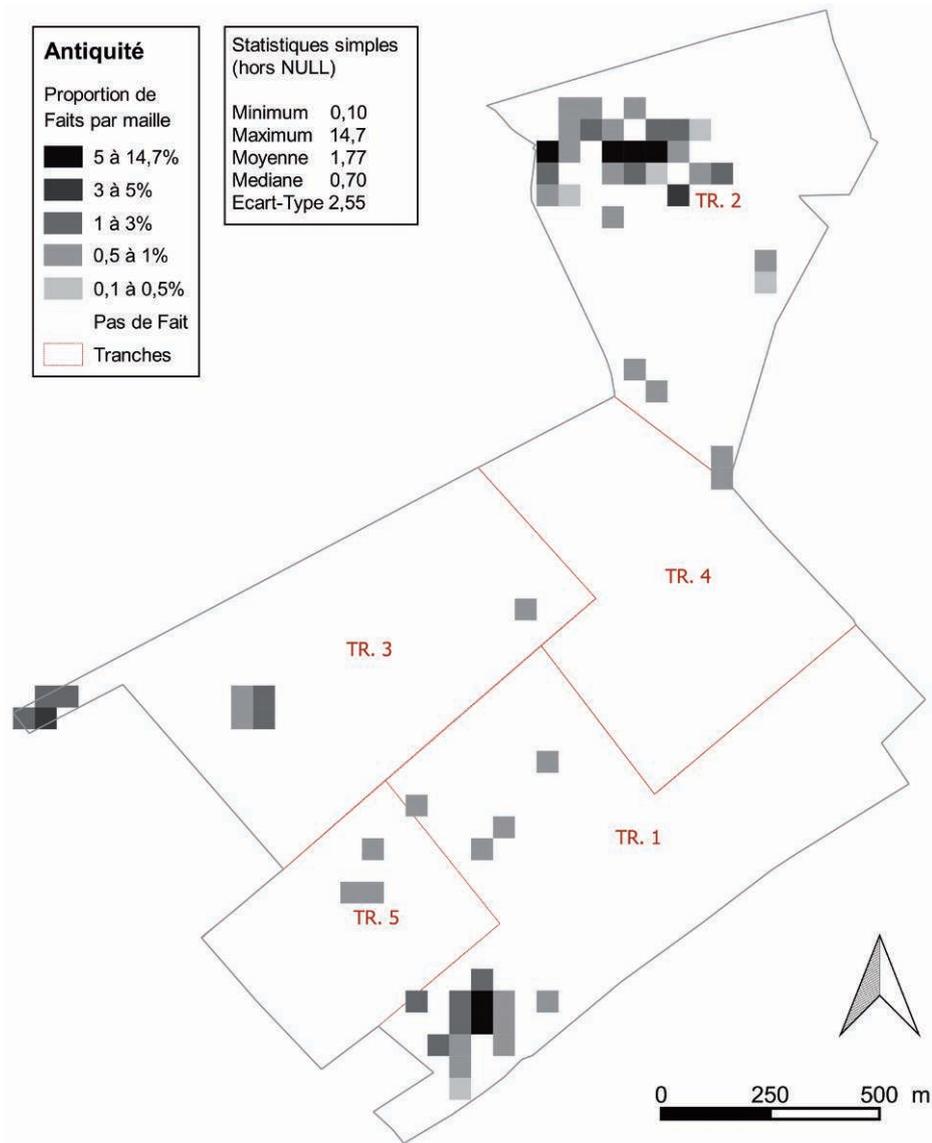


Fig. 7 : Analyse par maille de la proportion de Faits de l'Antiquité.

mais clairement associés à l'ensemble, essentiellement des trous de poteau et des fosses d'usages divers. Cette occupation bien délimitée s'étend chronologiquement du milieu du XI^e s. à l'époque Moderne sans discontinuités apparentes.

- Dans la tranche 2, la présence de Faits médiévaux est moins importante. En outre, la multiplicité des périodes présentes dans la même zone gêne la datation des Faits par association. Une concentration de Faits médiévaux suggérée par sept mailles à proportion moyenne à

forte se situe au centre de l'important ensemble antique décrit précédemment. Une occupation médiévale est ici superposée à l'occupation antique précédente et datée par une fourchette chronologique large, du VI^e s. au XI^e s.

La représentation par proportion de Faits au sein d'une maille par rapport à la totalité des Faits de la même période est facilitée ici par la présence de concentrations bien localisées. En revanche la faiblesse de l'échantillon, 74 Faits, rend la classification moins pertinente que pour les périodes mieux représentées. Un seul Fait daté, équivalant à plus de 1 %

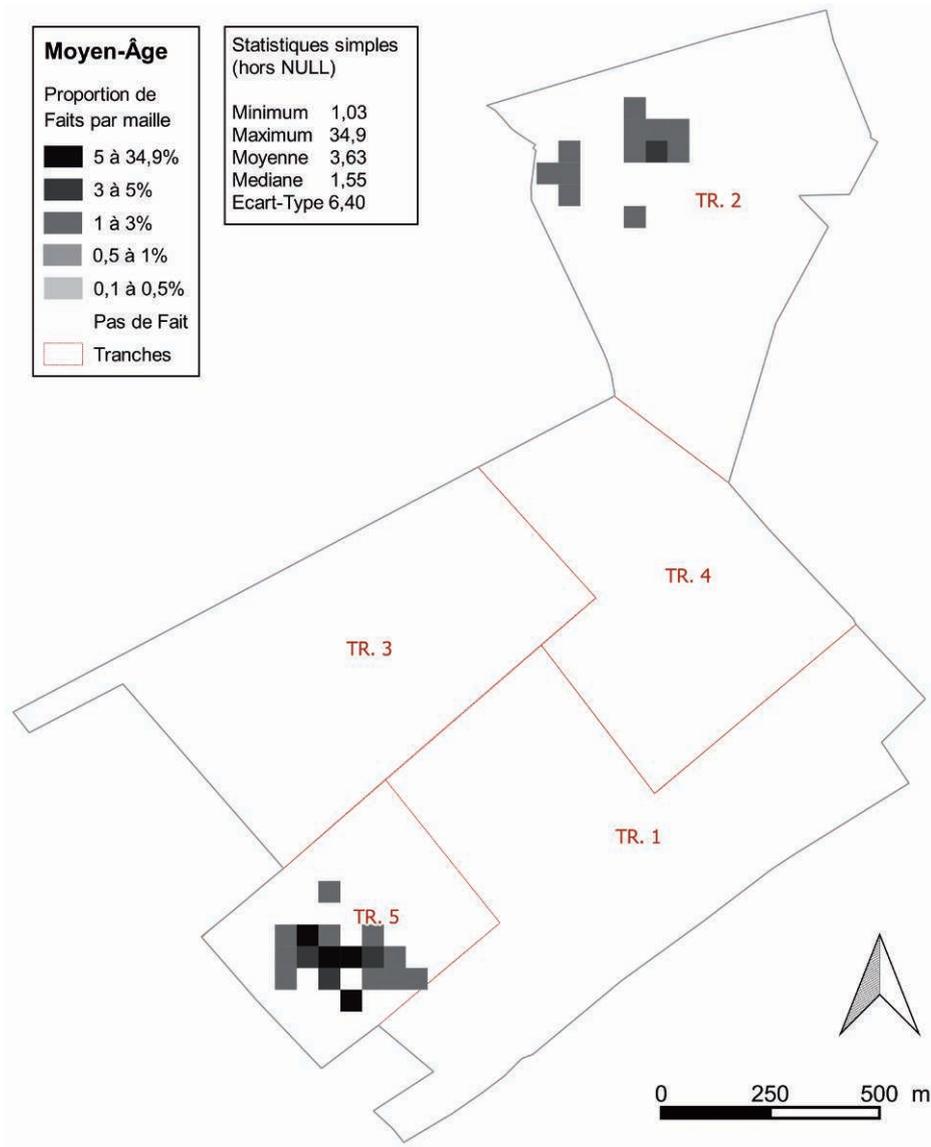


Fig. 8 : Analyse par maille de la proportion de Faits du Moyen Âge.

de la totalité des Faits, apparaît sous la forme d'une maille à proportion moyenne de 1 à 3 %. La période médiévale est donc bien représentée et interprétable à proximité de la seigneurie d'Ozans, au sud-ouest. En revanche, au nord, la présence humaine au Moyen Âge est plus difficile à caractériser. La comparaison entre les représentations des périodes deux à deux pourrait peut-être mettre en valeur les ruptures ou continuités de localisation des vestiges médiévaux aux phases antérieures et postérieures.

– L'époque Moderne (Fig. 9)

Cette période est bien représentée par 126 Faits datés, hors mobilier isolé, et exclusivement présente au sud-est du diagnostic sur la totalité de la tranche 5. Les vestiges de l'époque Moderne sont étroitement liés à ceux du Moyen Âge, la datation de certains Faits étant d'ailleurs difficilement dissociable. Un mur de clôture du domaine du château d'Ozans ainsi que des bâtiments maçonnés et des vestiges de sol du xv^e s. ou du xvi^e s. ont été découverts et sont représentés par les concentrations principales

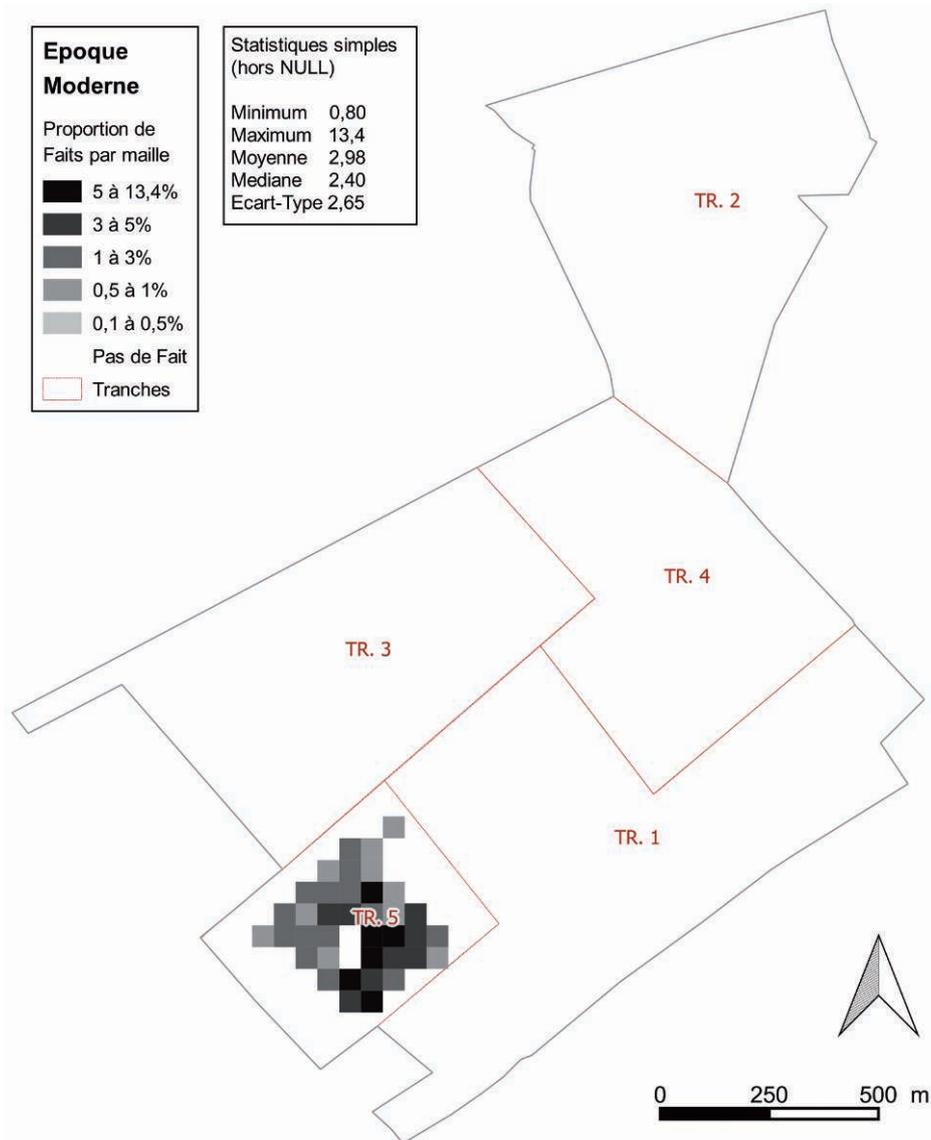


Fig. 9 : Analyse par maille de la proportion de Faits de l'époque Moderne.

de Faits dans les mailles aux valeurs les plus fortes. Des traces de parcellaire ainsi que de la route menant d'Ozans à Chignay et Renier au nord de l'ensemble sont encore visibles sur l'Atlas de Trudaine et parfois sur le cadastre napoléonien et complètent la représentation de l'occupation moderne.

Ici encore, la représentation par maille de la densité de vestiges est représentative de la distribution réelle des vestiges identifiés malgré l'imbrication des structures mises au jour avec celles de la période antérieure. La continuité ou les disparités de cette représentation d'une période à l'autre prend alors tout son intérêt.

1.6.2. Analyse par maille des différences de proportion de Faits entre deux périodes

Comme nous venons de le voir, la représentation par mailles carrées de 50 m de côté est pertinente pour la description et l'analyse des périodes étudiées à l'aide de la proportion de Faits par rapport à l'échantillon total de la même époque, pour peu que ce dernier soit assez important. Cette représentation est une généralisation de la distribution spatiale des Faits qui en permet la lecture, à l'échelle du diagnostic, en termes plus simples : présence/absence,

densité, organisation générale. Nous avons, sur la base de ces analyses par maille, produit des représentations des différences de proportions de Faits entre les périodes deux à deux, possibles uniquement si les deux échantillons sont quantitativement comparables. L'objectif est de caractériser les changements d'organisation généraux d'une époque à l'autre mais aussi, quand c'est possible, l'existence d'une continuité, d'un développement ou d'une différence dans les occupations d'une période à l'autre.

– *De la Protohistoire à l'Antiquité (Fig. 10)*

La représentation de la différence de densité de Faits par maille entre ces deux périodes montre avant tout la disparité de leur répartition dans la zone d'étude. En effet, alors que les Faits protohistoriques sont inégalement dispersés au sein de la zone d'étude, les vestiges antiques sont concentrés en agrégats bien distincts aux extrémités de l'emprise du diagnostic. Cette opposition est due à la différence de durée des deux périodes, la Protohistoire couvrant deux millénaires contre " seulement " cinq siècles pour l'Antiquité. L'essentiel des mailles à valeurs négatives, de -6,3 à -0,5, situées dans la zone centrale de l'emprise correspondent à la présence d'occupations protohistoriques ponctuelles localisés dans des espaces où la présence humaine durant l'Antiquité n'a pas été reconnue.

Au sud-ouest du diagnostic se mêlent une occupation protohistorique de l'âge du Bronze et une *villa* antique. Les valeurs négatives, en périphérie du site antique, indiquent le ou les emplacement(s) d'occupation(s) protohistorique(s) qui n'ont pas été occupées à la période antique ou dont le nombre de Faits est prédominant. Si l'on se rapproche du noyau antique, les valeurs ont tendances à s'annuler, de -0,5 à 0,5 %, montrant une superposition des Faits des deux époques. Enfin, au centre de cet ensemble, la proportion de Faits antiques est supérieure à très supérieure, de 0,5 à 14,7, à celle des Faits protohistoriques. Cette représentation montre ici la présence de deux occupations distinctes dans la même zone. L'apparente dispersion des vestiges protohistoriques serait alors due soit à l'organisation de l'occupation, soit à l'effacement des Faits les plus anciens par l'installation du site antique.

Au nord, les interprétations archéologiques n'ont pas permis de préciser l'organisation des vestiges protohistoriques auxquels succèdent deux sites antiques dont l'un s'étend principalement à l'extérieur de l'emprise. L'analyse de la densité de Faits de

chaque période et des différences de période à période accentue l'impression de dispersion des vestiges protohistoriques. Elle met aussi en évidence les mailles ne contenant qu'un Fait daté de la Protohistoire comme c'est le cas pour les cellules qui forment une ligne S/O-N/E traversant le site antique. Il convient donc de rester prudent afin de ne pas " surinterpréter " la carte. Au niveau de la seule concentration de Faits protohistoriques au nord de la zone, on n'observe pas d'extension du site antique.

La dispersion des vestiges protohistoriques par rapport à l'importante agrégation des Faits antiques n'est peut-être qu'apparente. En effet, le développement d'une occupation assez concentrée au cours de l'Antiquité a pu détruire les vestiges d'une occupation antérieure plus légère.

Enfin, la représentation de la différence entre les deux périodes a pour effet d'accentuer l'effet de concentration des vestiges sous la forme d'un groupe de cellules orangées à rouges qui apparaissent au sein d'un halo bleuté.

Ici, la méthode employée permet de mettre en évidence la dissemblance d'organisation spatiale entre les deux périodes. Par ailleurs, compte tenu de la différence d'ampleur chronologique entre les deux périodes, seule la datation de chaque concentration peut permettre d'émettre des hypothèses sur la continuité ou la discontinuité de l'occupation humaine.

– *De l'Antiquité au Moyen Âge (Fig. 11)*

La représentation des différences de proportion de Faits entre ces deux périodes est contrainte par l'inégalité des échantillons étudiés – 146 Faits antiques, 74 Faits médiévaux – mais cela est sans conséquence pour l'analyse lorsque les occupations ne se superposent pas.

La présence de Faits gallo-romains dispersés, en dehors des sites antiques déjà décrits, est ici mise en évidence par comparaison avec ceux de l'époque médiévale exclusivement concentrés au nord et au sud de l'emprise du diagnostic.

Au sud, la *villa* antique et les établissements médiévaux plus à l'ouest se distinguent spatialement et chronologiquement puisqu'un *hiatus* chronologique de plusieurs siècles les sépare, le premier datant des environs du III^e s. ap. J.-C., le second du XI^e s.

Au nord, la présence de vestiges des deux périodes sur la même zone se traduit par une prédominance des valeurs négatives, ce qui montre l'importance de la densité de Faits antiques par rapport aux Faits médiévaux. Les quelques mailles à valeurs

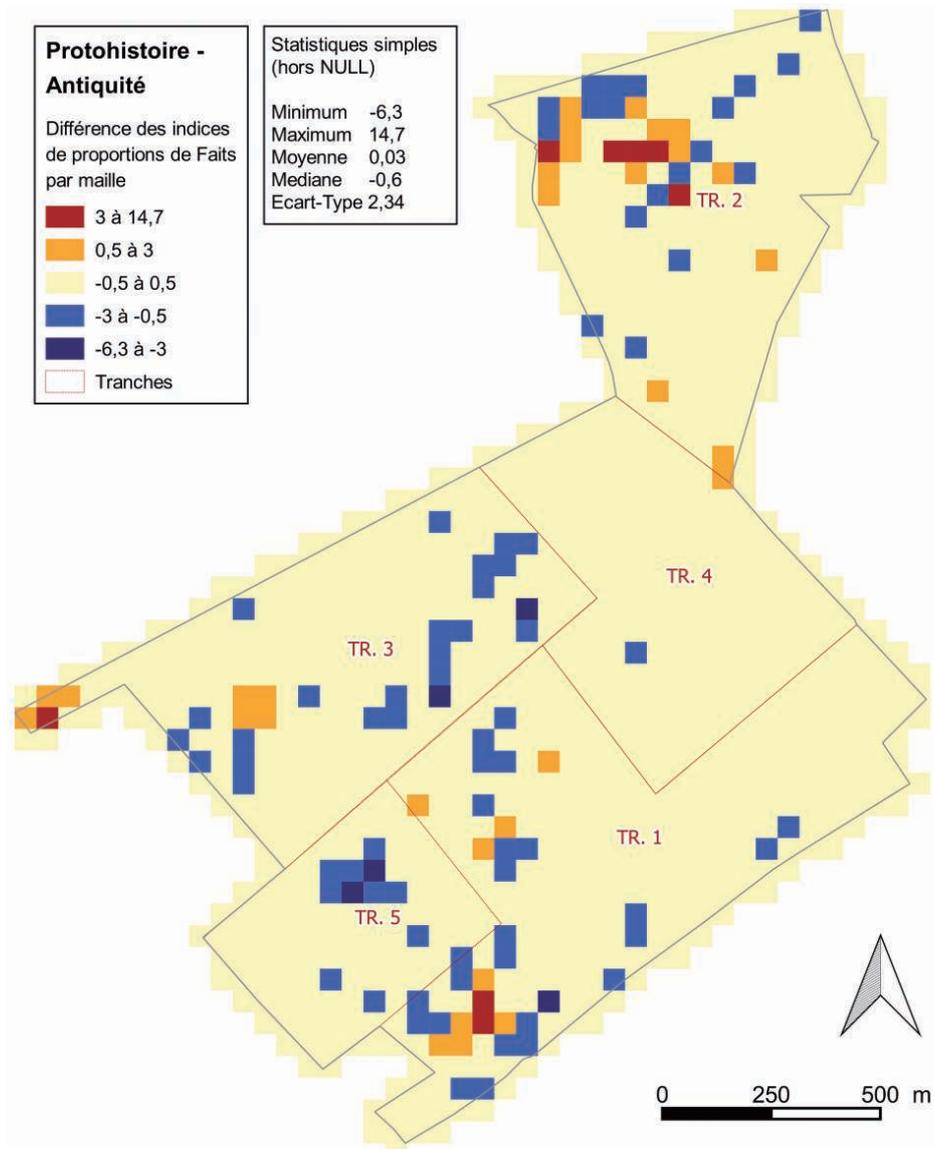


Fig. 10 : Analyse par maille des différences des indices de proportions de Faits de la Protohistoire et de l'Antiquité.

positives, de 0,5 à 3, que l'on pourrait interpréter comme la résultante d'une prépondérance des Faits médiévaux sont en fait dues à l'inégalité des échantillons. En effet, un Fait du Moyen Âge correspond à 1,4 % de la totalité des Faits médiévaux alors qu'un Fait antique représente moins de 0,1 % de l'échantillon ; la différence entre les deux sera donc forcément à l'avantage de la représentation des Faits médiévaux.

– Du Moyen Âge à l'époque Moderne (Fig. 12)

Pour ces deux périodes les occupations sont très concentrées, le Moyen Âge est représenté par deux pôles et la période moderne par une unique zone. De plus, ces deux époques sont étroitement liées à la seigneurie d'Ozans et son château situé au sud-ouest de l'emprise.

La différence de densité des Faits montre bien l'imbrication des deux périodes. Il y a d'une part un déplacement ou un agrandissement des installations

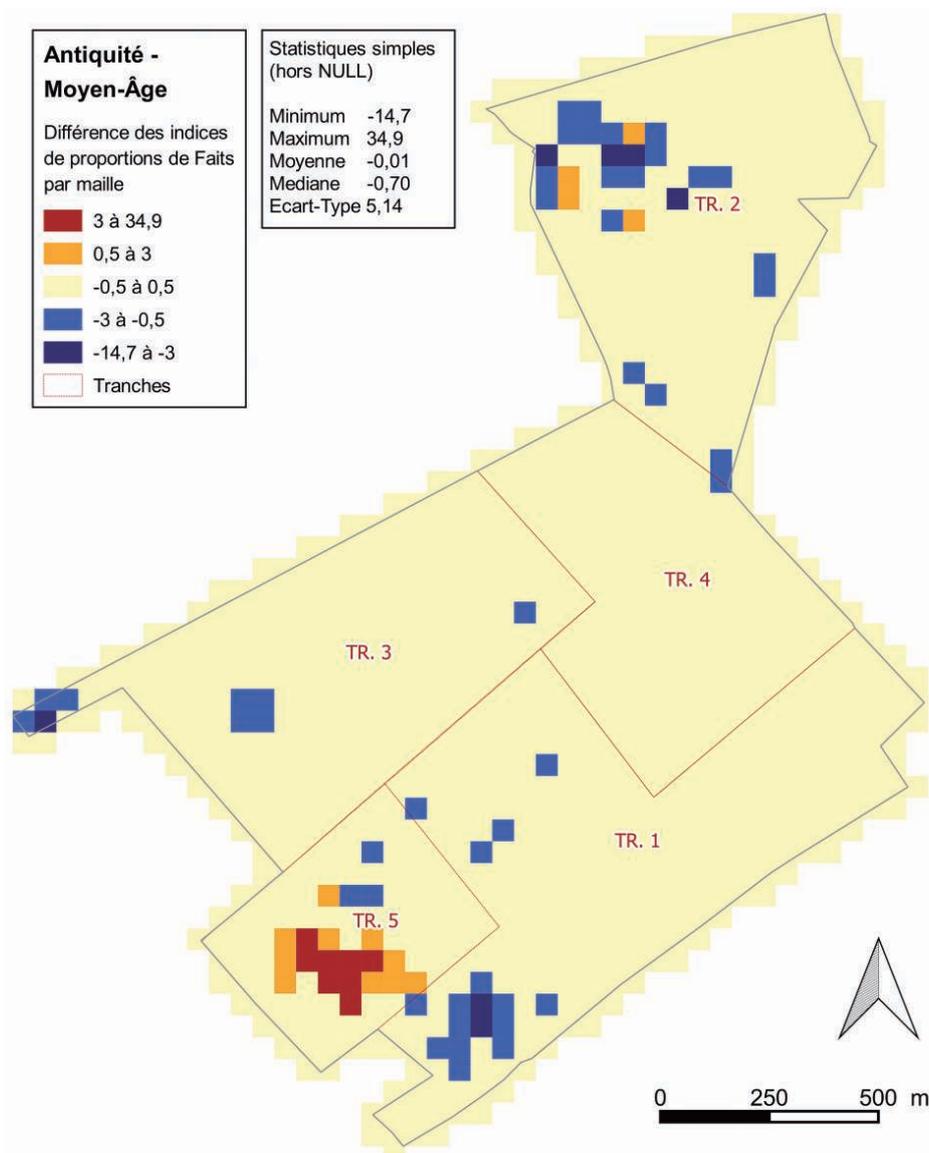


Fig. 11 : Analyse par maille des différences des indices de proportions de Faits de l'Antiquité et du Moyen Âge.

médiévales à l'époque Moderne, aux abords du château avec une différence d'indices de $-29,5$ à $-0,5$, d'autre part une forte densité de vestiges modernes est remarquable avec des valeurs positives, de $0,5$ à 3 , qui mettent en évidence les vestiges situés de part et d'autre de la route d'Ozans à Chignay sous la forme d'une ligne de carreaux S/O-N/E. L'expansion du domaine et de l'occupation à l'époque Moderne par rapport au Moyen Âge est représentée par des mailles avec des différences d'indice de $0,5$ à $10,3$ au nord-est de la zone. Finalement, une organisation spatiale se dessine à l'époque Moderne avec

deux orientations, S/O-N/E analogue à celle de la route, et sa perpendiculaire, S/E-N/O correspondant à celle du mur de clôture du château.

L'analyse des différences permet d'émettre des hypothèses concernant les changements d'occupation. Outre des différences de concentrations de Faits d'une période à l'autre, des continuités d'occupations ont été proposées ainsi que l'expansion de l'occupation humaine de la période médiévale à l'époque Moderne.

On peut conclure en précisant que l'analyse par maille révèle les concentrations de Faits par pé-

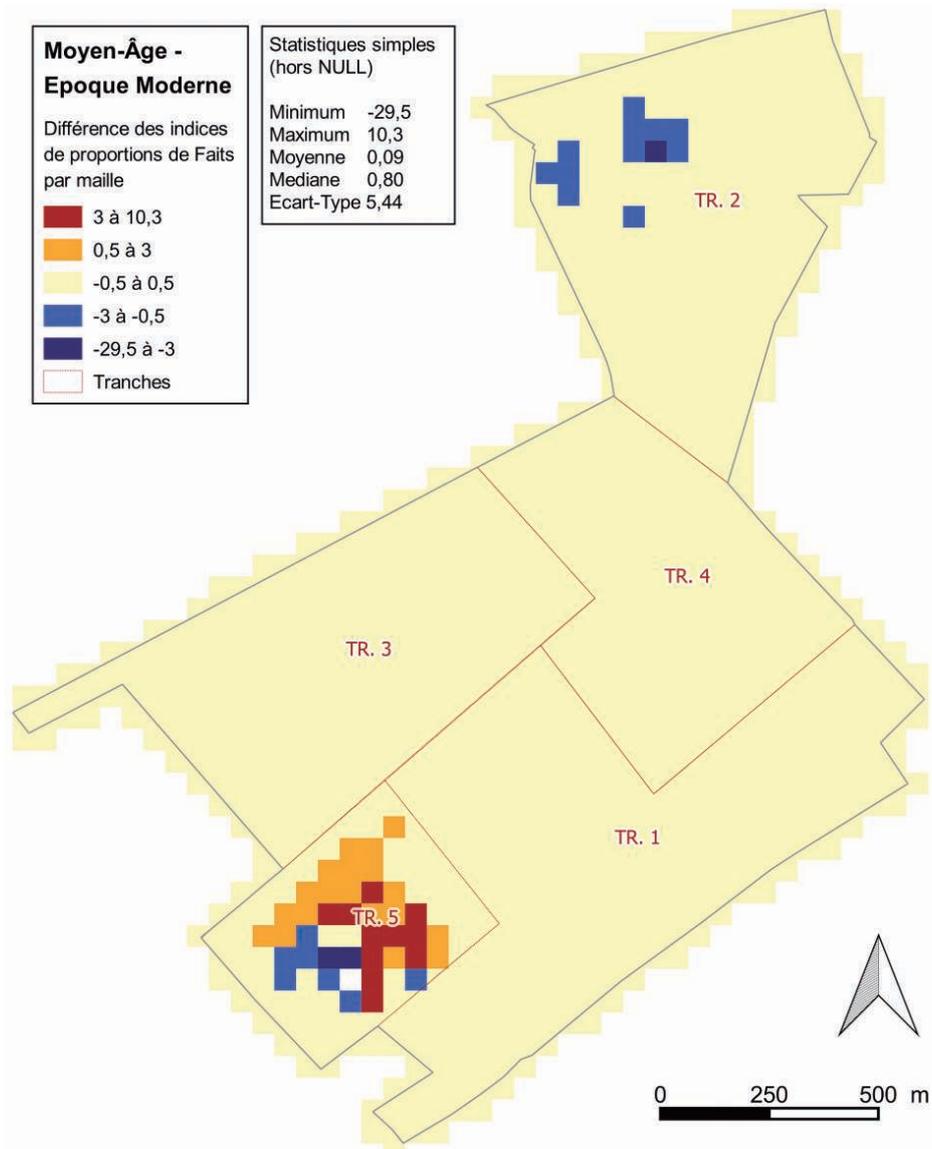


Fig. 12 : Analyse par maille des différences des indices de proportions de Faits du Moyen Âge et de l'époque Moderne.

riodes à l'échelle du diagnostic mais elle ne permet pas la localisation des concentrations de vestiges inférieures au grain de la maille ni la comparaison de toutes les périodes entre elles. L'analyse spatiale exploratoire est mobilisée en complément pour étudier la répartition par périodes des vestiges archéologiques sur la zone d'étude et émettre des hypothèses de continuités/discontinuités de l'occupation dans le temps.

2. ANALYSE SPATIALE EXPLORATOIRE

Nous considérons ici que l'espace géographique n'est pas un support neutre mais que l'espace influence les configurations naturelles et sociales (ZANINETTI 2005 ; PUMAIN et SAINT-JULIEN 2010). De même, l'étude de la distribution d'une information localisée implique l'acceptation de la première loi de la géographie telle quelle est définie par Waldo Tobler (TOBLER 1970 : 236) : " tout interagit avec tout, mais deux objets proches ont plus de chance

de le faire que deux objets éloignés ». En partant de ces postulats, nous cherchons à définir les propriétés géographiques qui régissent la répartition et l'organisation des Faits archéologiques d'une même période dans notre zone d'étude, puis à définir et à représenter des agrégats de Faits. Le protocole mis en œuvre comporte deux grandes étapes :

- L'utilisation d'analyses statistiques de premier et de second ordre afin de déterminer s'il est pertinent d'effectuer une recherche d'agrégats et, le cas échéant, de définir les échelles de concentration.
- L'application d'algorithmes de scan spatial afin de repérer les agrégats significatifs dans les semis de points.

Dans le cadre de cette analyse, les vestiges archéologiques, initialement représentés par des polygones, sont réduits à leurs centroïdes (centre de gravité de l'objet) pour obtenir un semis de points. Les analyses ont été réalisées avec les logiciels de statistiques spatiales CrimeStat III (LEVINE 2010) et de système d'information géographique QGIS (QGIS Development Team 2012).

2.1. La méthode du voisin le plus proche

Cette première méthode a pour but de répondre à deux questions :

- À l'échelle de la zone d'étude, la répartition des Faits archéologiques pour une période donnée peut-elle être assimilée à une distribution aléatoire, plutôt dispersée ou au contraire plutôt agrégée ?
- Si des concentrations ou agrégats sont présents, quelle est l'échelle optimale pour les observer ?

Elle s'appuie sur une propriété du premier ordre du semis de points : la distance moyenne d'un point quelconque à son voisin le plus proche. Cette distance est comparée à une distance théorique attendue sous hypothèse nulle de distribution aléatoire. Le résultat, la statistique R , ou ratio du voisin le plus proche est égal au rapport de la distance observée dans notre semis de points sur la distance attendue dans le cas d'une répartition aléatoire. Si $R = 1$, la distribution des points est aléatoire car la distance observée est égale à la distance attendue ; si $R = 0$, la distribution des points est agrégée car la distance observée est nulle et tous les points se touchent. Plus R tend vers 0 plus la répartition des points du semis est concentrée. Un ordre K est fixé afin de

s'intéresser à la distance au premier voisin puis au deuxième, puis au K^e .

Enfin, un test statistique permet d'évaluer la pertinence des résultats qui sont représentés sous la forme d'un graphique (Fig. 13).

2.2. Les résultats

Pour toutes les périodes étudiées, les semis de points ont une distribution relativement groupée avec une statistique R se rapprochant de 0. Toutes ont une p value² de 0,0001 c'est-à-dire que nous avons une chance sur 10 000 de conclure à la présence d'agrégats dans ces semis ponctuels alors qu'elle serait en réalité due au hasard.

Nous pouvons néanmoins distinguer des spécificités pour chaque période :

- Le semis de points des Faits protohistoriques contient des regroupements au niveau du premier voisin avec une statistique R de 0,444 puis se rapproche le plus d'une répartition aléatoire. Quand $K = 20$, elle est très proche d'une distribution aléatoire. Ce constat avait déjà été fait grâce à l'analyse par maille et trouve probablement son explication dans la durée de la période (deux millénaires), mais peut-être aussi dans l'organisation des sites de cette époque. La fonction indique une distance optimale de concentration entre 24 et 52 m pour l'ordre 1, c'est-à-dire pour agréger deux Faits.
- Le semis pour les Faits antiques, avec un R de 0,441, est aussi concentré que pour la période précédente. Le minimum de la fonction se situe à l'ordre 1 pour une distance optimale d'agrégation entre 25 et 62 m et un palier est atteint vers le sixième ordre, entre 80 et 168 m.
- Pour le Moyen Âge, le semis est très concentré avec une statistique R de 0,172. Cette forte concentration des vestiges avait déjà été mise en évidence lors de l'analyse par maille mais la méthode du voisin le plus proche d'ordre K apporte des informations supplémentaires. Le minimum de la fonction se situe à l'ordre 1 pour une distance optimale entre 15 et 87 m.

2. Dans un test statistique la p value définit la probabilité de commettre une erreur de première espèce, c'est-à-dire de rejeter à tort l'hypothèse nulle et donc d'obtenir un faux positif. Plus la valeur est faible, en général inférieure à 0,001, plus le résultat du test est statistiquement significatif.

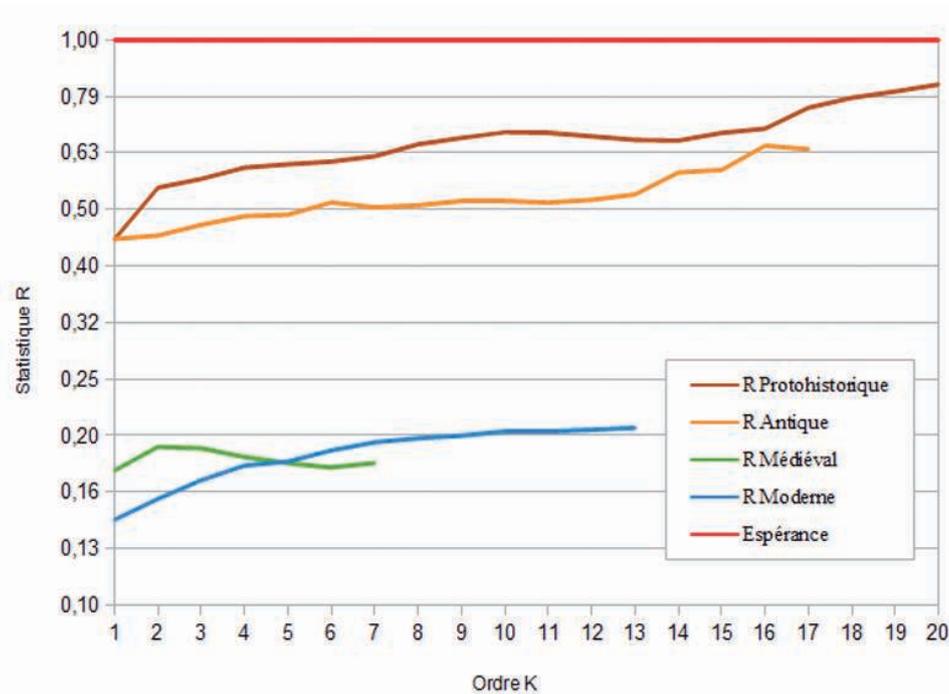


Fig. 13 : Comparaison de la statistique R pour les semis de points des Faits par périodes, d'après la méthode du voisin le plus proche d'ordre K (avec $K = 10\%$ de l'effectif).

Une seconde valeur faible de R (0,175) est atteinte au sixième ordre, elle correspond à une distance observée de 43 m et une distance attendue de 233 m entre lesquelles on peut observer de façon optimale les agrégats de six Faits.

- Enfin le semis de points modernes est celui dont la distribution est la plus groupée avec une très faible valeur de R (0,141), les agrégats étant situés entre 9 et 66 m. En outre, un palier se distingue aux alentours du cinquième ordre soit à une distance entre 29 et 162 m. Ces résultats sont la traduction statistique du regroupement des vestiges modernes dans une seule tranche du diagnostic.

Il y a donc nettement des concentrations spatiales pour toutes les périodes testées, avec des différences déjà pressenties avec l'analyse par maille voire parfois évidentes avec un simple plan général des vestiges. L'analyse permet néanmoins de fournir des outils d'analyse en définissant les échelles optimales d'observations des concentrations ainsi que le nombre de Faits qui forment un agrégat en fonction de l'échelle d'observation.

Pour tous les semis ponctuels, l'hypothèse nulle de répartition aléatoire est rejetée. Il convient ensuite de vérifier l'existence des agrégats localisés

et leur dimension moyenne, c'est l'objectif des analyses de second ordre et en particulier les fonctions K de Ripley et L de Besag.

2.3. Les fonctions K de Ripley et L de Besag

L'analyse spatiale exploratoire permet de définir le type de structuration spatiale auquel répond la distribution des Faits par périodes ainsi qu'une première approche des échelles d'agrégation et du nombre de Faits qui les composent grâce à la méthode du plus proche voisin d'ordre K. Les analyses de second ordre ont pour but, ici, de mettre en évidence les échelles d'organisation des Faits par périodes. Ces analyses de second ordre sont en mesure de répondre aux questions suivantes :

- L'espace est-il organisé en agrégats sources de dépendance spatiale ou les structures sont-elles réparties régulièrement ?
- Quelle est l'échelle de ces formes spatiales ?

Les analyses de second ordre ont été développées par B.D. Ripley (1976). Il s'intéressait à la probabilité d'observer un nombre donné de points dans un cercle autour d'un point quelconque du plan. Sous l'hypothèse de stationnarité, la densité à l'intérieur

du cercle devrait être égale à la densité de l'ensemble de la zone d'étude. La fonction K de Ripley compare cette densité attendue à la densité observée pour l'ensemble du semis étudié en traçant des cercles de recherches de rayon croissant. Les différents rayons de recherche produisent un corrélogramme graphique, représentant la fonction d'agrégation spatiale (en ordonnée) et la distance de recherche (en abscisse). On repère l'échelle d'organisation des agrégats dans les données par les points d'inflexion de la fonction.

La fonction K suit approximativement une loi exponentielle croissante, et son espérance est croissante. C'est pourquoi Besag (1977) en a proposé une transformée, plus facile à lire, d'espérance nulle. Ainsi la fonction L de Besag se prête à la comparaison de plusieurs zones ou de plusieurs semis de la même zone. L'interprétation de la fonction dérivée de Besag est plus facile. Les résultats sont représentés sous forme d'un graphique (diagramme cartésien) avec la distance en abscisse et la fonction L en ordonnées. Une correction des effets de bords est nécessaire afin que les points situés en bord d'emprise puissent être pris en compte au même titre que les autres. Par prudence, il convient de limiter la taille maximale du rayon de recherche à la moitié de la valeur du petit côté du rectangle englobant l'emprise. Les simulations réalisées produisent un intervalle de confiance sous hypothèse nulle de distribution aléatoire sous la forme de deux courbes (L_{min} et L_{max} , en traits fins sur la Fig. 14) qui encadrent l'espérance ($L = 0$ en noir sur la Fig. 14).

À l'échelle considérée, la courbe dessinée par la statistique L se lit ainsi :

- Lorsque L est comprise dans l'intervalle de confiance, alors la distribution observée est conforme à une distribution aléatoire homogène, dite Loi de Poisson.
- Lorsque L est plus élevée que la borne supérieure de l'intervalle de confiance, alors la distribution est concentrée.
- Lorsque L est plus faible que la borne inférieure de l'intervalle de confiance, alors la distribution est régulièrement dispersée.

L'évolution de la courbe selon les pas de distance permet de tirer des conclusions sur la structure spatiale générale. L'observation de changements d'échelle dans la structure spatiale du semis peut guider les étapes ultérieures de la recherche. " La fonction L permet la comparaison synchronique des structures spatiales de différents semis. Elle permet aussi de comparer différents semis sur le même territoire [...]. Une application particulièrement intéressante est d'interpréter la dynamique spatio-temporelle des

semis étudiés par l'analyse de semis successif sur le même espace d'étude " (ZANINETTI 2005).

Les résultats³ importés dans un tableur permettent de réaliser un diagramme de la fonction, le corrélogramme de Ripley, pour chaque période et d'étudier les valeurs, progressions et *maxima* de la fonction L. Enfin, un graphique récapitulatif a été réalisé pour comparer les différentes courbes des fonctions L (Fig. 14).

2.4. Les résultats

Le corrélogramme de Ripley (Besag) de chaque période permet une comparaison rapide de la distribution spatiale des différents semis de points. Contrairement à l'analyse de premier ordre, ici le semis de points de la Protohistoire se dégage nettement de ceux des autres périodes avec une tendance à l'agrégation beaucoup plus faible. Les semis des trois périodes suivantes présentent en revanche des similitudes, avec une tendance forte au regroupement jusqu'à atteindre un maximum puis décroître en s'approchant peu à peu de la convergence avec l'intervalle de confiance.

- La fonction L calculée pour le semis de points de la Protohistoire (Fig. 14, courbe marron) ne présente pas de pic net, elle est moins agrégée que celle des autres périodes. En fait, on distingue plusieurs échelles différentes d'agrégation : pour un rayon de recherche des agrégats de 15 m d'abord puis de 65 m et enfin avec une valeur maximale entre 280 m et 346 m. Ensuite l'intensité d'agrégation diminue jusqu'à se rapprocher de l'intervalle de confiance, sans jamais l'atteindre.

On observe donc une structure agrégée des faits protohistoriques à plusieurs échelles :

3. Plusieurs logiciels ont été testés sur un semis ponctuel. Nous avons retenu le logiciel CrimeStat car les paramétrages sont simples à mettre en œuvre et le graphique représentant les résultats d'une lecture plus aisée.

Après avoir défini l'étendue de la zone d'étude, sa surface et le système de coordonnées nous avons, après de nombreux essais, paramétré : une correction des effets de bords circulaires, un rayon de recherche à 501 m (qui respecte la règle de prudence qui veut que ce rayon soit au maximum égal à la moitié du petit côté du rectangle d'étude, 1 045 m dans notre cas). Enfin, des limites logicielles ont limité le test à 200 simulations (nombre suffisant de tests au vu du peu de différences entre les paliers 100, 150, 200 et 250 simulations).

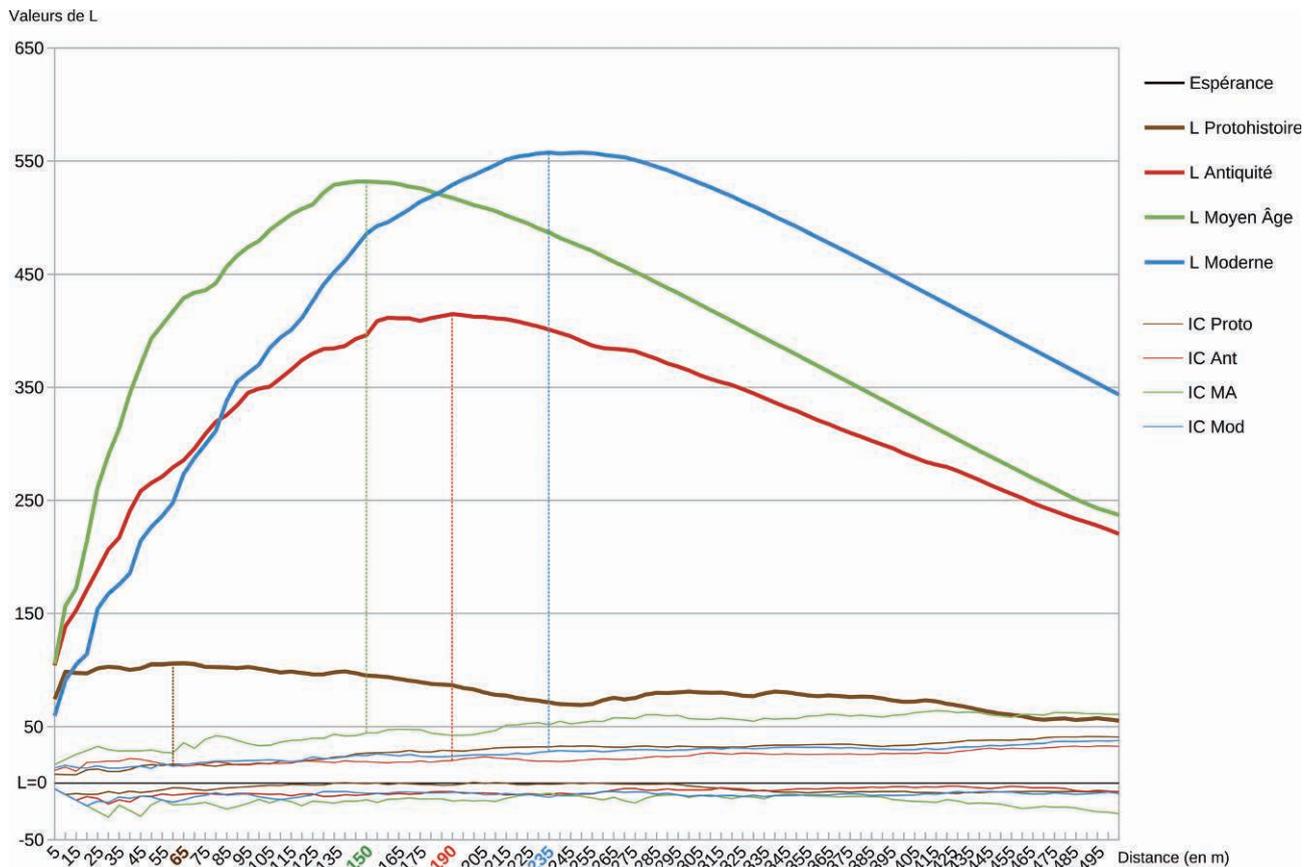


Fig. 14 : Corrélogramme de Ripley (Besag) pour les semis de points de Faits archéologiques découverts lors du diagnostic de la ZAC d'Ozans. Les courbes fines représentent les différents intervalles de confiance, les lignes en tirets indiquent les valeurs maximales de la fonction et la courbe noire l'espérance ($L = 0$).

d'une part dans un rayon de 15 et 65 m, ce qui confirme les résultats de la méthode du voisin le plus proche, d'autre part à une plus petite échelle dans un rayon de 280 à 350 m. Cette dernière pourrait correspondre à l'échelle d'observation optimale pour des ensembles de Faits cohérents. Comme nous l'avons déjà rappelé, les deux mille ans que couvre la période influent fortement sur la cohérence et la structure de l'organisation spatiale.

- Pour l'Antiquité (Fig. 14, courbe rouge), la fonction calculée a une forme de cloche avec une forte croissance jusqu'à un rayon de 45 m et une valeur maximale atteinte à 190 m. Alors que l'analyse de premier ordre semblait montrer une similitude avec la Protohistoire, la fonction L pour l'Antiquité se distingue clairement de celle de la période précédente

avec une structure beaucoup plus concentrée. Deux échelles d'observation des agrégats sont identifiées. Cependant, la présence de deux sites en limite d'emprises fausse la recherche d'échelles de concentrations malgré les corrections d'effets de bord.

- Pour le Moyen Âge (Fig. 14 courbe verte), la forme de la courbe est identique avec une intensité d'agrégation plus importante présentant deux inflexions et un palier entre des rayons de recherche compris entre 45 et 65 m. L'optimum de la fonction est à 150 m. Les résultats correspondent à la répartition des Faits que nous laissait entrevoir l'analyse par maille, avec une échelle idéale d'observation des agrégats de 150 m de côté, c'est-à-dire probablement au niveau des sites. À une échelle plus grande, déjà indiquée par la statistique R

aux alentours de 15 à 87 m, il existe un niveau optimal d'observation des agrégations, que la fonction L fixe autour de 45 m.

- La période ayant la distribution de Faits la plus concentrée est sans aucun doute l'époque Moderne, dont les vestiges se concentrent sur une seule tranche du diagnostic, et que la statistique R traduit par une valeur très faible (0,141). La fonction L (Fig. 7, courbe bleue) présente ici aussi une forme de cloche caractéristique des semis agrégés. Le corrélogramme pour cette période franchit deux inflexions à 90 m puis à 150 m avant d'atteindre un plateau, pour des distances de 200 à 280 m.

L'analyse de second ordre nous renseigne sur l'échelle optimale d'observation des agrégations aux alentours de 235 m mais aussi sur les échelles plus grandes, probablement dans les rayons de recherche de 90 m et de 150 m.

D'un point de vue général, la fonction L calculée pour les Faits des trois périodes les plus récentes présente une forme de cloche quasi similaire et représentative des semis agrégés (voir les exemples dans la thèse de F. Goreau⁴). On constate donc un phénomène d'agrégation important à grande échelle jusqu'à un optimum, puis une diminution des concentrations de Faits. La fonction L calculée pour la Protohistoire présente une courbe différente des autres avec une distribution des Faits moins agrégée sans pics net.

Ces analyses de la zone d'étude qui permettent de définir le type de structuration spatiale (agrégés ou pas) et les échelles optimales d'observation des agrégats sont le préalable à l'analyse locale et notamment à l'utilisation des algorithmes de scan spatial afin de repérer les agrégats significatifs dans les semis de points.

3. REPRÉSENTATION DES AGRÉGATS AVEC UN ALGORITHME DE SCAN SPATIAL : LA MÉTHODE DES FENÊTRES MOBILES

Un algorithme de scan spatial permet d'effectuer une analyse locale d'un semis de points, c'est-à-dire de déterminer des agrégats locaux significatifs à l'échelle de la zone d'étude. Le résultat de cette méthode peut être cartographié par des ellipses correspondant aux

4. Goreaud F. - Apports de l'analyse de la structure spatiale en forêt tempérée à l'étude et la modélisation des peuplements complexes, ENGREF (AgroParisTech), (2000 : 59 et 99).

agrégats. Les algorithmes de scan spatial sont comparables à la fonction K de Ripley (qui compare la densité à l'intérieur d'un cercle à celle qu'elle devrait être dans l'hypothèse nulle de stationnarité) mais en faisant une analyse ligne par ligne, point par point. Le logiciel renvoie alors les fenêtres de concentration supérieures à ce que l'on pouvait attendre dans le cas d'une hypothèse de stationnarité. Cette algorithme de scan spatial est appelé " méthode des fenêtres mobiles " car il identifie les agrégats à l'intérieur d'une fenêtre circulaire qui balaye la zone d'étude⁵. Cette méthode permet de tester des variantes en changeant le rayon de recherche et le nombre d'individus minimums dans un agrégat après avoir :

- vérifié l'utilité de la démarche par les méthodes du voisin le plus proche et K de Ripley ;
- optimisé le rayon de recherche sous l'hypothèse de stationnarité (ordre K du voisin le plus proche et corrélogramme de Ripley).

Le rayon de recherche, noté h , est obtenu par les résultats précédents avec la fonction K de Ripley (Besag) en prenant les valeurs maximales, c'est-à-dire lorsque la courbe est la plus éloignée de l'intervalle de confiance (Fig. 14). La dimension minimale des agrégats selon Zaninetti (2005) doit être supérieure au produit de la surface du cercle de rayon h , qui servira de fenêtre mobile, et de la densité du semis. Les résultats de ce calcul sont présentés dans un tableau récapitulatif (Fig. 15).

Les résultats

Les résultats du scan spatial sont représentés sous forme d'ellipses de déviations autour du barycentre des agrégats identifiés ou de polygones. Le scan spatial a été réalisé pour chaque période puis les résultats ont été assemblés sur une seule carte (Fig. 16).

5. C'est ici encore avec le logiciel CrimeStat que cette méthode a été mise en œuvre, avec le module intitulé STAC (*Spatial and Temporal Analysis of Crime*). Le logiciel nous permet de définir le rayon des cercles qui serviront de fenêtres mobiles de scan, mais aussi le nombre de points dans cette fenêtre qui forment un agrégat.

Lors de la mise en œuvre de l'analyse, il faut veiller à éviter deux écueils :

- Le défaut de précision : définir un cercle de scan de taille trop importante ou un trop petit nombre d'individus minimum pour former un agrégat.
- Le défaut de puissance : demander un nombre important de points pour former un agrégat, ce qui réduira le nombre d'agrégats retenus au final.

Périodes	Protohistoire	Antiquité	Moyen Âge	Ep. Moderne
Maximum de la fonction $K = \text{rayon de recherche} = h$	65	190	150	235
Surface du cercle de la fenêtre mobile (de rayon h)	13266,5	113354	70650	173406,5
Nombre d'individus total = N	205	146	76	130
Densité du semis = $N/\text{surface totale}$	0,000090815	0,000064678	0,000033668	0,00005759
Nombre <i>minimum</i> d'individus formant un agrégat	1,2	7,3	2,4	10,0

Fig. 15 : Calculs du rayon de recherche et de la dimension minimale des agrégats pour le paramétrage du scan spatial (méthode des fenêtres mobiles).



Fig. 16 : Résultats du scan spatial dit " méthode des fenêtres mobiles " sous forme d'ellipses de déviations, par périodes. r = rayon de recherche de la fenêtre mobile et ind. = nombre d'individu minimal pour former un agrégat.

- Les ensembles retenus par le scan spatial pour la Protohistoire (fenêtre mobile circulaire de 65 m de rayon et deux individus *minimum* par agrégats) sont assez proches des interprétations issues de l'analyse par maille avec la dispersion évidente des vestiges de cette période. Au sud, seuls deux ensembles ont été retenus à l'est et à l'ouest de la *villa*.

Plus au nord, près de la cour du château, l'occupation est bien délimitée, l'importance du nombre de Faits et leur éloignement des autres vestiges de la même période ayant facilité la tâche.

Au centre, parmi la distribution apparemment éclatée des vestiges, le scan spatial met en évidence cinq agrégats de petite taille qui correspondent aux sites déjà décrits mais dont la représentation est ici mieux localisée.

Enfin, au nord, la Protohistoire est représentée par deux concentrations distinctes, au nord et au centre du site antique.

À cette échelle, le scan spatial n'apporte que peu d'informations supplémentaires par rapport à l'analyse par maille, la dispersion générale et l'étendue de l'intervalle chronologique y étant sûrement pour beaucoup.

- La méthode appliquée pour l'Antiquité, avec un rayon de recherche de 190 m et un minimum de huit Faits par agrégat, permet de distinguer les deux plus importants sites repérés ainsi qu'une autre concentration de Faits.

La *villa*, au sud de l'emprise est représentée par une ellipse de déviation prenant en compte la plupart des Faits, preuve d'une concentration importante à l'échelle d'observation. Le niveau d'observation des agrégats pour cette période semble ici pertinent.

Les résultats sont moins probants à l'extrême nord de l'emprise ou la concentration de Faits retenus par la méthode donne une ellipse qui englobe presque tous les vestiges. Alors que l'analyse par maille permettait de distinguer un site partiellement contenu dans l'emprise à l'ouest, le scan spatial regroupe les deux occupations sans discrimination. La nature incomplète de l'échantillon montre une des limites de la méthode statistique censée être effectuée sur un échantillon exhaustif.

- Enfin un dernier agrégat se trouve à l'ouest où une concentration de Faits isolés a été interprétée comme un site dont l'étendue dépasse les limites de l'emprise.

L'échelle d'analyse semble correspondre aux établissements humains de cette période mais la méthode montre ses limites face à un échantillon peu fiable.

- Les Faits du Moyen Âge, dont la distribution a été caractérisée comme très concentrée, ont fait l'objet d'un scan spatial avec un rayon de 150 m pour la fenêtre mobile et un minimum de trois individus par agrégats.

L'ellipse résultant du scan spatial n'a pris en compte que les concentrations les plus importantes en délimitant ainsi le cœur du site médiéval mis au jour.

Au nord, alors que l'analyse par maille semblait démontrer l'existence de deux agrégations de Faits, le scan n'en retient qu'une représentée par l'ellipse de déviation.

Pour la période médiévale, l'échelle d'analyse obtenue par les analyses semble pertinente même si un probable site partiellement inclus dans la zone d'étude, à l'extrême nord-ouest, se trouve éliminé.

- Enfin la méthode appliquée aux Faits de la période moderne, avec un rayon de recherche de 235 m et une dimension minimale d'agrégations de dix Faits, prend en compte la totalité des vestiges de la période et dessine une ellipse cohérente avec l'interprétation de l'analyse par maille. Ce résultat est sans surprise étant donné la concentration des Faits dans une seule tranche du diagnostic, et il confirme les précédentes conclusions sur la période.

L'utilisation d'un algorithme de scan spatial, avec la méthode des fenêtres mobiles et les étapes préalables indispensables pour s'assurer de la pertinence de la démarche, des échelles d'observations et de la dimension des agrégats, est donc probante même si elle nécessite d'en connaître les limites et les facteurs influant sur le résultat, notamment la qualité de l'échantillon. Cette méthode démontre tout son intérêt quand on affiche les résultats des différentes analyses sur une même carte (Fig. 16).

En effet, contrairement à l'analyse par maille qui ne permettait que la comparaison des concentrations de Faits pour deux périodes, les ellipses issues du scan spatial peuvent toutes être représentées ensemble et rendre la comparaison des agrégats plus aisée pour toutes les périodes, à l'échelle de la zone d'étude. Cette superposition ne permet pas cependant d'interroger la continuité ou la discontinuité des occupations d'une période à une autre.

La modélisation de la répartition des Faits par périodes est un bon moyen de dégager les grandes

tendances de distribution des vestiges. Les niveaux de concentrations sont bien caractérisés depuis la dispersion des occupations protohistoriques jusqu'à l'extrême concentration de l'ensemble moderne. De même la multiplicité des périodes représentées sur une même zone est perceptible et leur assemblage peut être approché : l'*imbroglio* chronologique de la zone nord est plus lisible grâce à cette discrimination par périodes des agrégats.

La méthode des fenêtres mobiles est réputée être efficace pour repérer les agrégats d'échelle intermédiaire dans des semis relativement dispersés. Cette méthode relativement flexible, puisque l'utilisateur peut faire varier les paramètres (rayon de recherche et dimension des agrégats), souffre de l'inexistence de véritable test d'aide à la décision pour le paramétrage. L'application des paramètres issus de la méthode du plus proche voisin et de la fonction de Ripley (Besag) a cependant démontré ici son efficacité. La force de cette méthode réside, en fin de compte, dans l'aspect visuel des résultats.

CONCLUSION

Le résultat de cette étude synthétisé dans les cartes de superposition des ellipses de déviation issues du scan spatial et de l'analyse par maille (Fig. 17) offre une lecture synthétique de l'information issue du diagnostic fondée sur un protocole d'analyse spatiale robuste et reproductible. Il répond à l'objectif général : donner du sens aux données d'un diagnostic sur une grande surface grâce à des documents interprétables par le responsable d'opération et les prescripteurs.

Les méthodes utilisées permettent de s'affranchir du masque créé par l'implantation des tranchées et de représenter sur une carte des résultats pertinents de l'organisation générale des vestiges par périodes, pour peu que l'on maîtrise le processus analytique et les limites de la méthode. Les documents produits offrent une représentation plus lisible que les plans généraux des vestiges fournis dans les rapports mais qui ne rendent pas compte des concentrations révélatrices des occupations. Le responsable d'opération dispose ici d'une aide à l'interprétation archéologique, qui sans s'y substituer, renforce sa démarche hypothético-déductive grâce à une lecture synthétique de la masse importante de données produite par un diagnostic archéologique de plusieurs centaines d'hectares.

Les cartes de répartition des vestiges archéologiques, par périodes et entre elles, sont une géné-

ralisation de l'information qui prend néanmoins en compte tous les faits archéologiques datés quelle que soit leur position dans l'emprise étudiée. En effet, le caractère simplificateur des résultats est fondé sur des protocoles d'analyses stricts et un traitement systématique des données. Les documents produits rendent lisibles et comparables les données, qu'elles soient trop denses ou éparpillées. Ils peuvent, dans certains cas, aider à la compréhension de structures spatiales, voire en révéler. *A minima*, ils permettent de proposer des zones de concentration fondées sur l'intégralité du corpus et moins empiriques que celles tracées manuellement.

Ces méthodes d'analyse, fondées sur des outils mathématiques et statistiques, traitent la localisation des vestiges comme une de leur propriété intrinsèque. Elles ne peuvent être mises en œuvre qu'à la condition d'une formalisation de l'enregistrement et au géoréférencement des données de terrain pensées dès en amont de l'opération.

Finalement, la mise en œuvre de ses méthodes d'analyse spatiale et les documents qui en résultent servent aux responsables d'opération dans leur démarche analytique et leur argumentation mais peuvent également servir d'outils d'aide à la décision pour les prescripteurs.

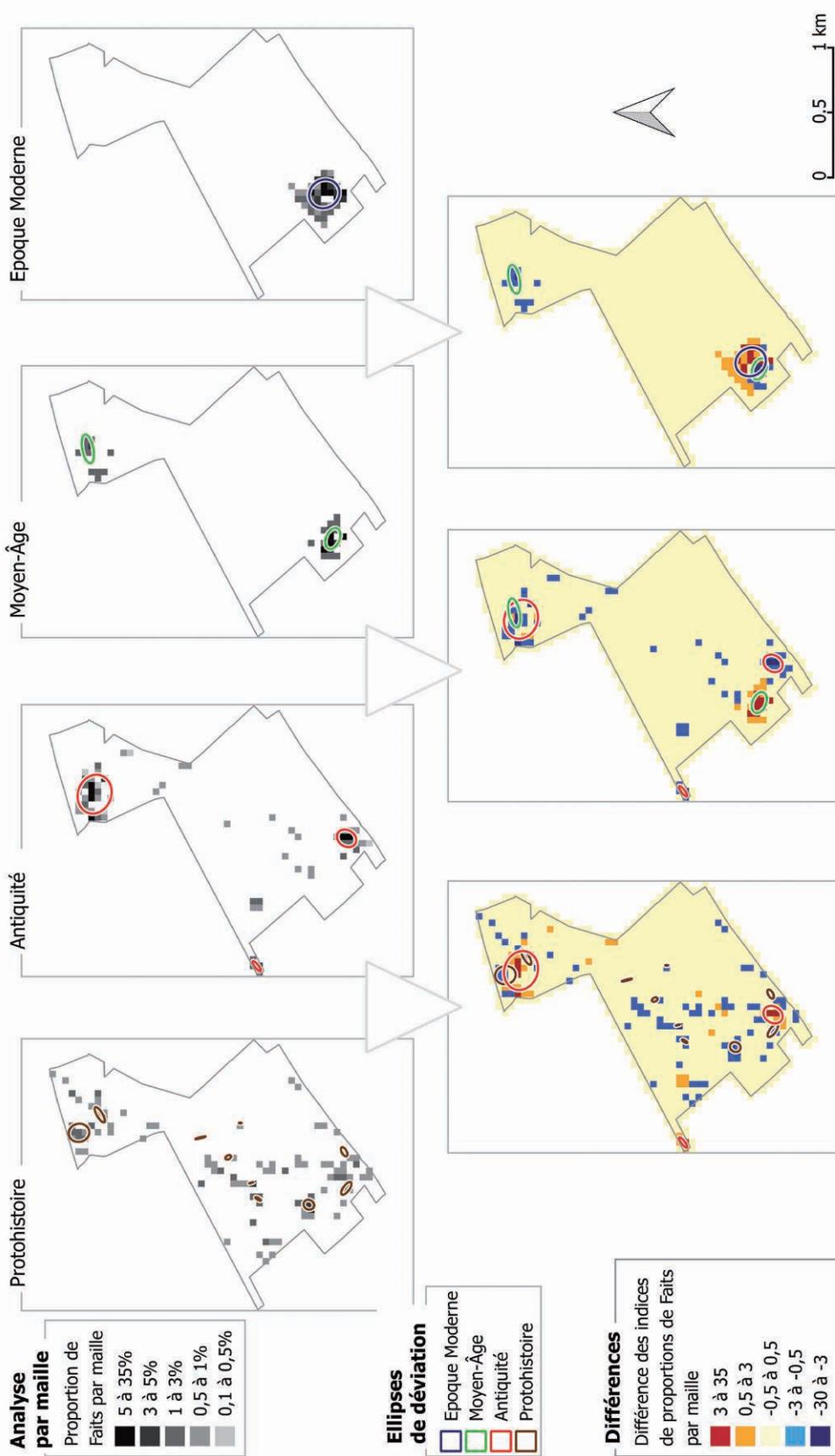


Fig 17 : Superposition de l'analyse par maille (proportion de Faits par périodes et différences) et des ellipses de déviations issues d'un algorithme de scan spatial.

BIBLIOGRAPHIE

BESAG 1977

Besag J. E. - Contribution to the discussion of Dr Ripley's paper, *Journal of the Royal Statistical Society B.*, 39 : 139-195.

BLANQUAERT et MEDLYCOTT 2006

Blancquaert G. et Medlycott M. - *Archeological Evaluation of Rural Area of North West Europe*, Planarch, Maidstone, 42 p.

BRUN et VANMOERKERKE 2006

Brun P. et Vanmoerkerke J. - Essai d'évaluation des opérations de grande surface, *Les nouvelles de l'archéologie* : 88-95.

DUBOULOZ 2003

Dubouloz J. - L'évaluation des méthodes de diagnostic : simulations sur des sites de l'Aisne, *Nouvelles de l'archéologie* : 46-50.

HOLZEM 2011

Holzem N. - *ZAC d'Ozans - Phase 1, Tranche 5 - " Le parc du château d'Ozans "*, Rapport de diagnostic archéologique, Inrap, Tours, 232 p.

LAJOIE 1991

Lajoie G. - L'analyse intra-urbaine par carroyage, *Geographica Polonica*, 61 : 47- 68.

LAJOIE et BUSSI 1993

Lajoie G. et Bussi M. - *Mise au point d'une cartographie par carroyage de l'espace urbain : contraintes techniques et validité statistique*, in : *Actes du 8^e colloque de Géographie théorique et quantitative*, Budapest.

LEVINE 2010

Levine N. - *CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations (v 3.3)*, Washington DC, Ned Levine & Associates, Houston (Texas) and The National institute of Justice.

MUNOS 2010a

Munos M. - *ZAC D'OZANS - Phase 1, Tranche 1*, Rapport de diagnostic archéologique, Inrap, Tours, 368 p.

MUNOS 2010b

Munos M. - *ZAC D'OZANS Phase 1, Tranche 4*, Rapport de diagnostic archéologique, Inrap, Tours, 173 p.

MUNOS 2010c

Munos M. - *ZAC D'OZANS - Phase 1, Tranche 3*, Rapport de diagnostic archéologique, Inrap, Tours, 214 p.

MUNOS 2010d

Munos M. - *ZAC D'OZANS - Phase 1, Tranche 2*, Rapport de diagnostic archéologique, Inrap, Tours, 338 p.

PUMAIN et SAINT-JULIEN 2010

Pumain D. et Saint-Julien T. - *L'analyse spatiale*, A. Colin, Paris.

QGIS DEVELOPMENT TEAM 2012

QGIS Development Team - *QGIS Geographic Information System. Open Source Geospatial Foundation Project*.

RIPLEY 1976

Ripley B. D. - The second order analysis of stationnary point process, *Journal of Applied Probability*, 13 : 255-266.

TOBLER 1970

Tobler W. - A Computer Movie Simulating Urban Growth in the Detroit Region, *Economic Geography*, 46 : 234-240.

VERHAGEN et BORSBOOM 2009

Verhagen P. et Borsboom A. - The design of effective and efficient trial trenching strategies for discovering archaeological sites, *Journal of Archaeological Science*, 36 : 1807-1815.

ZANINETTI 2005

Zaninetti J.-M. - *Statistique spatiale : méthodes et applications géomatiques*, Hermès science publications, 320 p.

