



HAL
open science

The Hesbaye and Condroz Regions (Belgium). Analysis of archaeological Site Patterns from Roman to Merovingian Times through logistic Regression Modelling

Hanne de Brue, Gert Verstraeten, Bastiaan Notebaert, Annick Lepot, Amélie Vallée

► To cite this version:

Hanne de Brue, Gert Verstraeten, Bastiaan Notebaert, Annick Lepot, Amélie Vallée. The Hesbaye and Condroz Regions (Belgium). Analysis of archaeological Site Patterns from Roman to Merovingian Times through logistic Regression Modelling. Michel Reddé. Gallia Rustica 1. Les campagnes du nord-est de la Gaule, de la fin de l'âge du Fer à l'Antiquité tardive, 1 (49), Ausonius éditions, pp.153-178, 2017, Mémoires, 978-2-35613-206-2. hal-03029565

HAL Id: hal-03029565

<https://hal.science/hal-03029565>

Submitted on 28 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

The Hesbaye and Condroz Regions (Belgium)

Analysis of archaeological Site Patterns from Roman to Merovingian Times through logistic Regression Modelling

Hanne De Brue, Gert Verstraeten, Bastiaan Notebaert, Annick Lepot and Amélie Vallée

Archaeological site patterns contain essential information to gain more insight in the interaction of past societies with the physical environment. This study investigates the spatial patterns of rural habitation and burial sites dating from the Early Roman, Late Roman and Merovingian period for two contrasting regions in central Belgium (i.e., the Hesbaye and Condroz). Rare events logistic regression analyses were applied to evaluate the impact of multiple environmental factors and the ancient road network on the presence of archaeological sites, and to map probabilities of site occurrence. Results indicate that burial sites tend to be located in river valleys, whereas settlements are more often situated at relatively elevated parts of the landscape. The regional site pattern is furthermore significantly affected by the proximity of loess and limestone in the Hesbaye and the Condroz, respectively. Diachronic analyses do not indicate major shifts in locational factors through time. Finally, comparison of observed site densities with the sensitivity for soil erosion and the presence of agricultural land use demonstrated that the archaeological record is significantly biased by the spatially differential preservation and discovery potential of sites.

INTRODUCTION

The invaluable source of information covered by the remains of past human occupation phases and their spatial distribution in the landscape holds the key to improve our understanding of various aspects of ancient civilisations' geographic behaviour. These include the societies' interactions with the physical environment as well as the temporal and spatial variation in the location preferences characteristic to different types of archaeological sites. Numerous studies have indeed described and analysed the spatial organisation of past and present settlement systems at various scales and in a variety of environments. Such analyses can either be performed in a theoretical-conceptual framework, e.g. by utilising cognitive models built on assumptions about individual and communal interactions such as economic efficiency or subsistence (e.g. Sanders 1956; Green 1963; Ellison & Harriss 1972; Binford 1982; Peterson & Drennan 2005), or alternatively through the application of empirical-inductive approaches to establish relations between the observed spatial distribution of archaeological sites or artefacts and one or more cultural or environmental factors. Early explorative work in the latter field by, e.g. Childe (1934), Fox (1943) and Grimes (1945), has laid the foundation for more recent developments in predictive modelling research, which explicitly aims to mathematically predict the (probability of) presence or the density of archaeological sites in poorly and unexplored geographic areas (Kohler & Parker 1986; Rose & Altschul 1988; Conolly & Lake 2006; Verhagen *et al.* 2010). Logistic regression analysis has shown to be a particularly suitable and robust technique to estimate site occurrence probabilities based on a set of spatial predictor variables that can be either categorical or numerical in nature and do not require normal distributions (e.g. Kvamme 1985; Parker 1985; Carmichael 1990; Warren & Asch 2000; Woodman 2000). Other methods that have been applied to

investigate location choices based on archaeological data sets include Monte Carlo simulations (e.g. Vanacker *et al.* 2001; Crema *et al.* 2010), Bayesian statistics (e.g. Buck *et al.* 1996; Verhagen 2006; Finke *et al.* 2008) and Dempster-Shafer theory (e.g. Canning 2005; Ejstrud 2005; van Leusen *et al.* 2009). An advantage of the latter two techniques is the possibility to determine spatial patterns of uncertainty on the estimated site occurrence probabilities. However, these approaches also require substantial prior knowledge, either from expert judgement or from other, purely empirical site analyses.

Variables reflecting former natural resources that have possibly affected archaeological site patterns include topographic features, as well as the soil's fertility, drainage and erodibility, native vegetation biomes, and the proximity of streams and coasts. Undoubtedly, a wide range of ideological, political, social and economic factors have played a decisive role in the location choices of past societies. The integration of the often hardly quantifiable and subtle effects of cultural variables in predictive modelling studies of archaeological site patterns remains hitherto challenging (Kohler 1988; Conolly & Lake 2006). Nevertheless, empirical relations that are primarily or even solely based on environmental parameters, by critics often referred to as ecological determinism (e.g. Butzer 1982; Wheatley 1992; Ebert 2000), may provide valuable insights into the spatial behaviour of cultures as well. As a matter of fact, the application of deterministic models allows in itself to assess if and to what extent site patterns are, directly or indirectly, explained by physical landscape properties.

Inevitably, the preservation of archaeological remains, and hence also any relation describing their spatial distribution, is affected by numerous post-depositional natural and anthropogenic processes. Moreover, both contemporary land use and the intensity and methods of archaeological prospection lead to a spatial differentiation in the sites' discovery potential. Although the consequential incompleteness of the archaeological record induces a bias in predictive model outcomes, such models also allow to evaluate the influence of variables reflecting preservation and discovery processes (e.g. soil erosion, prospection quality, land use) on observed site patterns.

In this study, rare events multivariate logistic regression analyses were applied to a database of Early Roman to Merovingian rural burial and habitation sites for two contrasting regions in central Belgium (i.e., the Hesbaye and Condroz) in order to map probabilities of archaeological site occurrence, and to assess the impact of the ancient road network and the topographic, hydrological, lithological and soil characteristics of the landscape on the observed site patterns. More generally, the presented study aims to evaluate whether the locations of a limited number of archaeological sites can be used to accurately quantify probabilities of site occurrence as a proxy for human activity, and as such improve our understanding of the spatial and temporal evolution of small-scale occupation patterns and the extent to which the latter are determined by the physical properties of the environment. In addition, the preservation and discovery bias in the spatial distribution of archaeological sites is investigated by comparing observed site densities with the soil's sensitivity for erosion and with the presence of agricultural land use.

STUDY AREA

Physical environment

Factors determining the spatial pattern of Roman and Merovingian archaeological sites were investigated in two contrasting Belgian natural environments, i.e., the regions of the Hesbaye and Condroz (fig. 1). In this study, the easternmost part of the Brabant Plateau (i.e., the Dijle catchment) was included in the Hesbaye region. Also note that the studied regions were delineated based on modern municipal boundaries to facilitate the collection of archaeological data, and may therefore deviate slightly from the strictly physical-geographical boundaries.

The Hesbaye area covers *ca.* 3630 km² and is characterised by a rolling topography. Altitudes reach approximately 250 m a.s.l. in the south, with an average of *ca.* 100 m, while river valleys are incised down to *ca.* 15 m. In the southern part of the region, called the "Dry Hesbaye", a thick layer of Quaternary aeolian loess deposits covers the permeable substrate comprised mainly of Tertiary sands and Cretaceous chalk (fig. 2), creating a low-density river network with numerous dry valleys, and highly fertile, well-drained loamy soils (Maréchal & Ameryckx 1992). The northern part of the Hesbaye ("Humid Hesbaye") and the Brabant Plateau in the west are marked by a more shallow loess to sandy loess cover, with local outcrops of Tertiary sands and clays. The less permeable substrate causes a more pronounced relief in this area compared to the Dry Hesbaye, with a dense

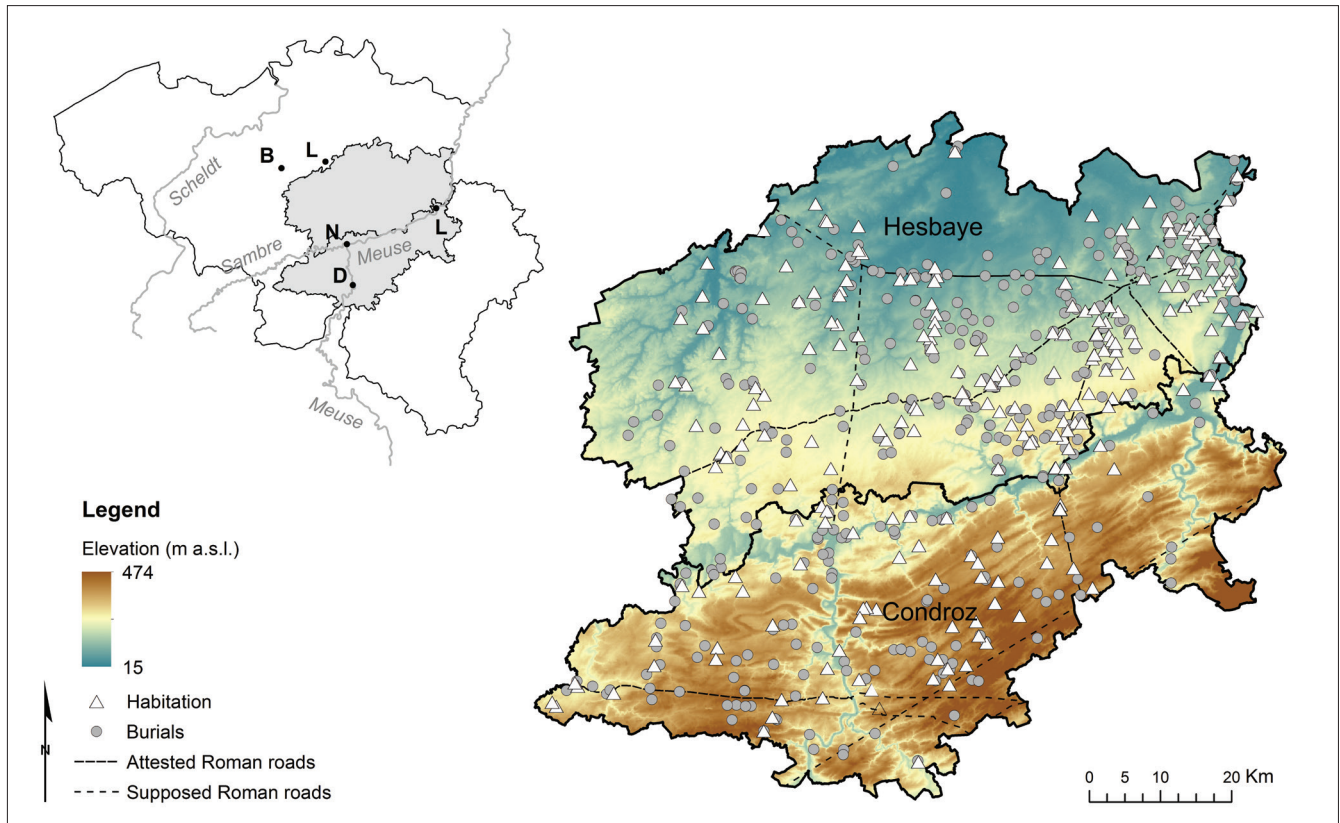


Fig. 1. Elevation and archaeological site locations dating from the Early Roman to Merovingian period in the Hesbaye and Condroz regions. The location of the study area within Belgium is marked on the inset map, together with major contemporary cities (in clockwise order, B: Brussels; L: Leuven; L: Liège; D: Dinant; N: Namur).

network of rather deeply incised river valleys (De Moor & Pissart 1992). Although the fertile loess deposits are less thick and on hilltops sometimes absent, the area remains highly suitable for crop cultivation. The modern rural landscape indeed has an open-field character with vast areas under agriculture and only small forested patches.

The Condroz region (*ca.* 2720 km²) is situated south of the deeply incised valleys of the Sambre and Meuse rivers, and north of the Fagne and Famenne depressions. The parallel alternation of Carboniferous psammite ridges and upper Devonian limestone depressions (fig. 2), oriented west-east or southwest-northeast, is characteristic for the region and creates an undulating relief with an average elevation of *ca.* 220 m a.s.l. (De Moor & Pissart 1992). The northernmost part of the Condroz has a more smooth, elevated topography shaped by resistant lower Devonian and Silurian rocks. Large parts of the Palaeozoic substrate of the Condroz are covered with loess, resulting in loamy soils that can be very shallow and often contain psammite or limestone gravelly inclusions (Maréchal & Ameryckx 1992). Rivers in the Condroz generally follow the geological structures, resulting in a lattice of rather small rivers located in limestone depressions and several larger, deeply incised transversely oriented streams such as the Meuse, Ourthe and Hoyoux (De Moor & Pissart 1992). Today, the landscape is dominated by agriculture on the loess plateaus and pasture in the valleys, whereas the shallow, stony soils on ridges and steep slopes are covered with deciduous and coniferous forests.

Annual precipitation in the study area varies from less than 700 mm in the northern Hesbaye, to *ca.* 900 mm in the most elevated parts of the Condroz, and is spread more or less equally throughout the year (Dupriez & Sneyers 1978. See also: Brulard 1960; Godart & Feltz 2009; Godart & Feltz 2010).

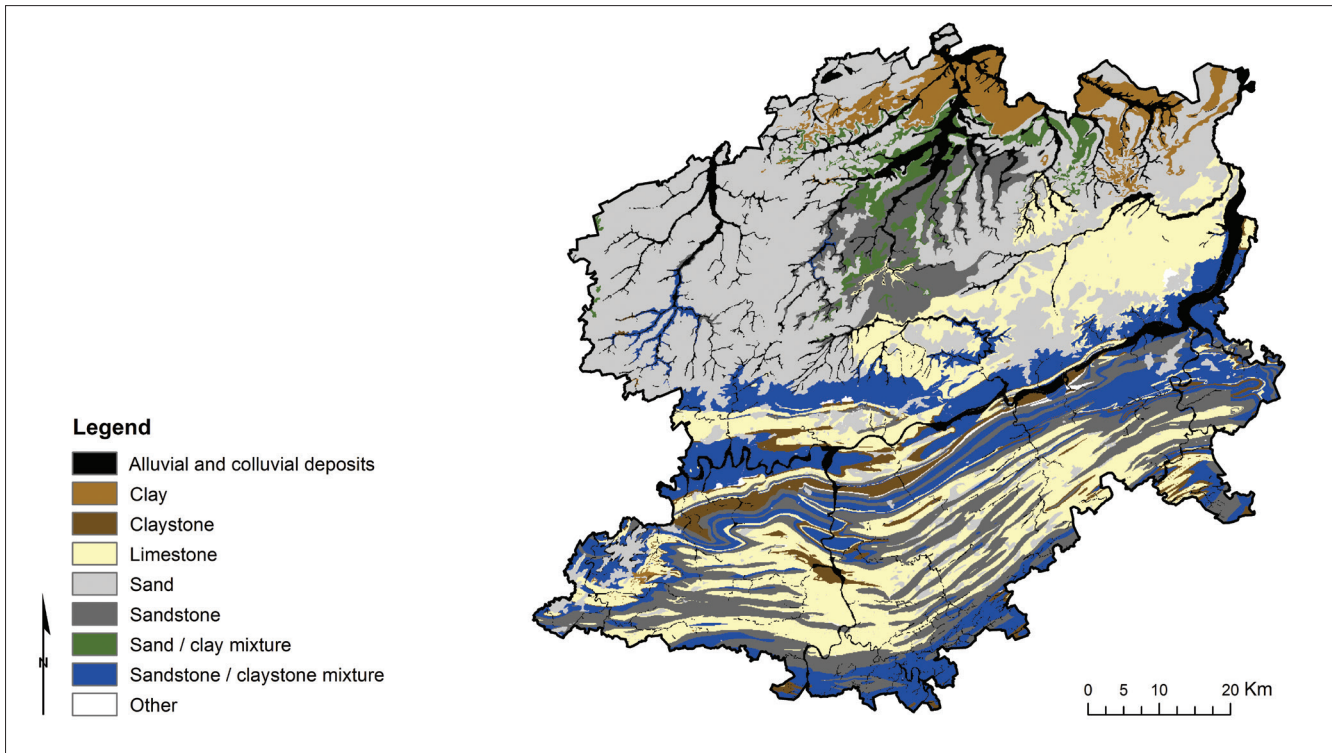


Fig. 2. Simplified lithology of the Hesbaye and Condroz regions, based on the geological maps of Belgium (1:40,000; © Geological Survey of Belgium, 2007). More detailed lithological types (e.g. schist, chalk and marble, psammite and quartzite) were included in the general categories claystone, limestone and sandstone.

Historical and archaeological context

The rich history of rural human occupation in the Belgian loess area during Roman (*ca.* 50 BC - AD 450) and Merovingian times (until AD 751) is attested by numerous studies (fig. 1; regional overviews can be found in, e.g., de Boe 1971; Raepsaet & Raepsaet-Charlier 1975; Van Ossel 1992; Nouwen 2006; Brulet 2008a; Verslype 2013), as well as palynological (Broothaerts *et al.* 2014) and geomorphological (Vanwalleghem *et al.* 2006; Rommens *et al.* 2007; Notebaert *et al.* 2011) proxy records.

The actual Hesbaye and Condroz regions were located in the central part of the *civitas Tungrorum* (fig. 3), probably set up by Drusus between 12 and 9 BC (Raepsaet-Charlier & Raepsaet 2011). Whereas its first belonging to the province of *Gallia Belgica* or *Germania* is still discussed, the *civitas* was definitely integrated to the *Germania inferior* during Domitian's reign in 85 AD (Raepsaet 2013). The earliest Roman settlements in Tongeren, the *caput civitas*, developed along the Bavay-Tongeren road, possibly around 10 BC (Vanderhoeven 1996, 221). However, most of the agglomerations evolved more significantly along the roads and the Meuse River from the middle of the 1st century AD onwards, after the development of the Roman communication network (Brulet 2008a, 109-110). At the same time in the countryside, the Early Roman rural landscape of the Hesbaye and Condroz was scattered with both indigenous wooden byre-houses and stone-built villas (Nouwen 2006; Brulet 2008a). Whereas the villas likely adopted a cereal-based monoculture (combined with small-scale gardening) generating a surplus for nearby rural centres, the persistence of byre-houses indicates that pastoral activities still remained, even in the most fertile regions of the *civitas Tungrorum* (e.g. Roymans 1996; Brulet 2008a; De Clercq 2009; Pigièrè 2015).

Due to the recession of the 3rd century, socio-economic issues induced a military and political crisis which developed significantly through the second half of the century (Brulet 2008a; De Clercq 2009). As part of the implementation of the dioceses, Tongeren joined the province of *Germania II* at least by the end of the 3rd century (Opsteijn & Taayke 1998; Vanderhoeven 2002; Nouwen 2006). The city kept its capital status despite a decrease of the urban space surrounded by a new but shorter town wall. During the Late Roman Period, most of the road agglomerations declined rapidly before their complete abandonment,

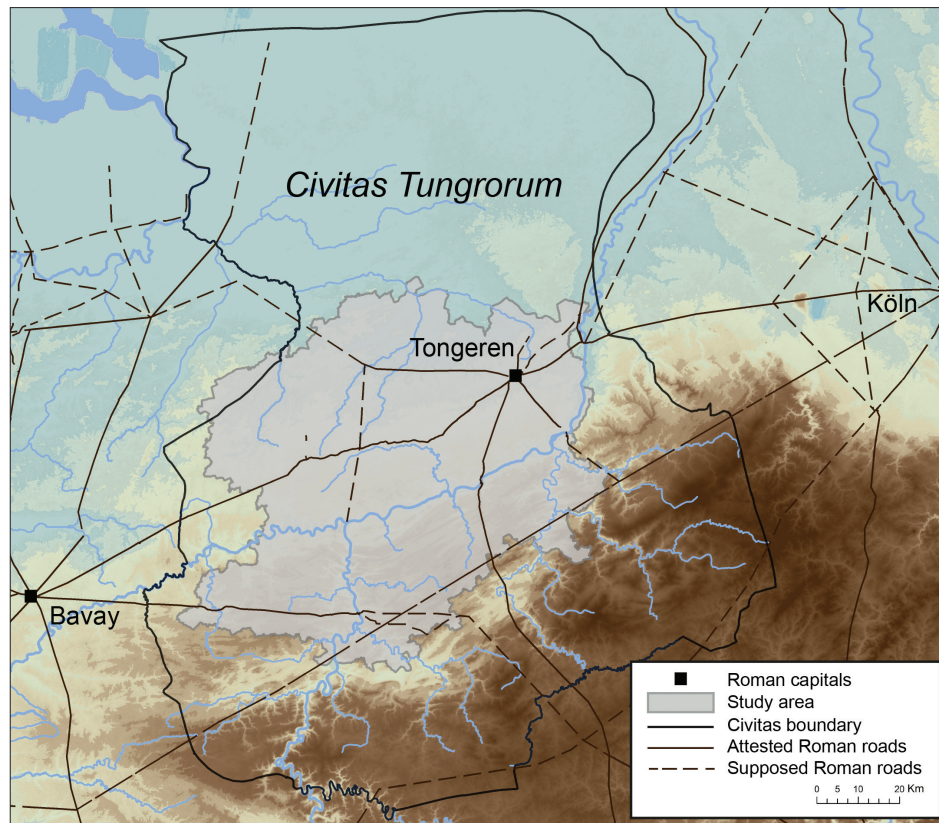


Fig. 3. Location of the study area within the roman boundary of the *Civitas Tungrorum*.

while those lying along the course of the Meuse survived. An inner defensive system, including fortifications sometimes built in the midst and upon the ruins of the former agglomerations, was simultaneously set up along the Bavay-Köln road. After the mid 3rd century, a series of fortifications was also created on rocky headlands overhanging the Meuse River and its tributaries (Brulet 2008b). Some Early Roman rural settlements were however still occupied in the 4th century, and sometimes reoccupied by wooden farmhouses settled next to such villas as in Wange or Neerharen-Rekem (Van Ossel 1992; Raepsaet-Charlier & Vanderhoeven 2004; Van Ossel 2010). That period shows the transition from a specialised economy to a more decentralised and self-sufficient system (Bitter 1991), as well as a general decrease in open, cultivated space and husbandry practices (Defgnée & Munaut 1996a; 1996b ; Kooistra *et al.* 2004 ; Pigière 2009).

At the end of the 5th century, the region was integrated in the *Regnum Francorum* by Clovis Ist herited by his son Thierry Ist in 511, at the western edge of the Austrasian kingdom (Bührer-Thierry & Mériaux 2010, 129, 139). The bishop's see, established in Tongeren since the 4th century, moved to Maastricht at least in the 6th century (Gauthier 2002). The Merovingian economic renewal is attested by the increase of several production centres spotted along the Meuse from the middle of the 5th century onwards. Evidences of rural occupation mainly fall within funerary archaeology (e.g. Roosens 1949 ; Heymans 1978; Verslype 2003a; Verslype 2013). However, the discovery of scarce rural settlements suggests that wooden byre-houses occurred dispersed (Rogge 1981; Hollevoet 1991; Verslype 2010; Hollevoet 2016). The growing influence of the aristocratic family in the land management is clearer in the 7th century. The Pippinids, owning many land proprieties in middle Meuse basin, would soon push this region, for a time, in the core of the Carolingian empire.

The evolution of the natural environment in northern Gaul was the subject of several studies on the territory of actual Belgium (Defgnée & Munaut 1996; Kooistra 1996; Matteredne 2001; Broothaerts *et al.* 2014). Deforestation began long before the Roman conquest. In the eastern part of the loess belt, i.e. Hesbaye region, the Late La Tène landscape was already widely open.

Pollen analysis reveals a more extensive use of fields and pastures and an increase in cereal pollen (Munaut 1967; Kooistra 1996). This phenomenon increased over two centuries during the Roman period. Cattle husbandry became of major economic importance, which can be partly explained by an intensification of agriculture in the area and the need of cattle for traction. The agricultural growth in this zone during the Roman period is illustrated by changes in land use, choice of crops (mainly spelt), technical equipment, and social organisation of the production (Roymans 1996, 58-65). At the same time, the breeding of cattle responded to the need for mass production of meat and raw material for craft activities in urban centres (Pigière 2009; Vanderhoeven & Ervynck 2007). It is only at the beginning of the 3rd century AD and especially in the 4th century that pollen analysis allows to identify a reforestation in the loess belt (Kooistra *et al.* 2004). This reflects a possible change in land use that probably influenced farming practices or the reverse. In the western part of Hesbaye, pollen data show that deforestation mainly began from the beginning of the Bronze Age onwards and reached highest levels during the Early Roman Period. Reforestation is observed at the beginning of 2nd century AD. This period with decreased human impact on the vegetation lasted until the beginning of the 5th century AD (Broothaerts *et al.* 2014). These overall observations are however far from materialising the variety of situations that could coexist depending on the micro-regions, the location of settlements or the agrarian system put in place by the inhabitants within each site. The methods adopted to face such bias are explained in the next section.

METHODS

The archaeological data

The archaeological records come from an extensive database built in the framework of the Belspo IAP VI/22 and CORES VII/09 projects. Archaeological sites were spatially digitised by systematically checking all available sources and records in the study area. They can be subdivided in the database according to their age, interpretation, spatial location, etc. (see also Lepot *et al.* 2016). A four-level precision scale allows to deal with the degree of accuracy of each site location (for instance recent excavations *vs* sources from the 19th historical and archaeological societies). Each region was treated considering the specificity and the quality of the available sources, aiming at the recording of the archaeological data. Some parts of the study area with a short archaeological prospection history, as well as areas with poor site visibility or access (e.g. forested, urbanised or industrial zones), may remain blank and are therefore not as representative. It should also be noted that only well-documented, -interpretable and -datable archaeological structures were recorded. Superficially prospected sites (resulting from pedestrian surveys for instance) and small or isolated finds (e.g. coins, potsherds), difficult to date or interpret, were not included in this inventory.

For the logistic regression analyses, the raw data available for the Hesbaye and Condroz regions were refined and classified. We chose to work only with rural settlements and funerary sites dating from the 1st to the 7th century. Indeed, urbanised settlements and late antique fortified hilltop settlements are assumed to reflect a specific set of location strategies and were therefore excluded from the analyses. Moreover, we only took into account the site locations with a sufficiently high spatial accuracy, i.e., up to a few 100 metres.

All the sites are dating back to between 50 BC and AD 751. They are classified into Early Roman (50 BC to AD 270), Late Roman (AD 270 to AD 450) and Merovingian periods (AD 450 to AD 751). The Early Roman time is the best documented one with 481 occurrences. Settlements include large and small villas as well as some traditional timber-built farmhouses and annexes. Funerary sites encompass cemeteries, isolated graves and tumuli (typical individual monumental graves). The Late Roman documented settlements do not exceed the number of 59. These are selected because they are revealing a sustaining occupation during this time period. 126 Merovingian sites are also registered. The best documented type among them are the cemeteries whereas habitation sites gather only some wooden buildings, pit-houses, but mainly dispersed pits of which the fillings include pottery sherds. Persisting settlements were assigned to all relevant periods. The majority of these selected sites may be localised with an accuracy of less than 100 metres.

Rare events logistic regression analyses

The relation between a dichotomous response variable Y , i.e., presence (1) or absence (0) of an archaeological site, and a set of numerical and/or categorical explanatory variables x_1, x_2, \dots, x_n , can be described by the logistic function

$$P(Y=1) = \frac{1}{1+e^{-(\hat{\alpha}+\hat{\beta}_1x_1+\hat{\beta}_2x_2+\dots+\hat{\beta}_nx_n)}} \quad (1)$$

where $P(Y=1)$ is the probability of site occurrence, and $\hat{\alpha}, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_n$ are the maximum-likelihood regression coefficients. However, as the presence of an archaeological site can be considered a rare event, meaning that the total number of site observations in the area of interest is many times smaller than the area without archaeological evidence, application of ordinary logistic regression could lead to sharply underestimated site occurrence probabilities. King and Zeng (2001b) proposed a set of rare events corrections for logistic models, including endogenous or case-control stratified sampling, an intercept correction for selection bias taking into account prior knowledge on the true fraction of events in the population, and a correction for the parameter estimates that deals with the bias related to finite statistical samples, which is amplified by rare events. Their methods are implemented in the Zelig package (Imai *et al.* 2011), available for R software (R Core Team, 2013). Rare events logistic regression has been applied to deal with the scarcity of positive observations in a variety of studies, including landslide susceptibility assessment (Van Den Eeckhaut *et al.* 2006; Bai *et al.* 2010), the formation of gullies and closed depressions (Vanwallegem *et al.* 2008), and factors influencing international political conflicts (King & Zeng 2001a).

In this study, rare events logistic regression analyses were applied in order to determine the relation between the occurrence of various combinations of archaeological site type, period and region as the dependent variable (tabl. 1), and multiple explanatory variables (tabl. 2). All raster maps of the independent variables were processed at a spatial resolution of 100 m. Aspect class, landform, lithology, soil drainage and texture, and the occurrence of agricultural land use were treated as categorical or dummy variables, while all other explanatory variables are numerical. Note that the distance to the main Roman roads is the only explanatory variable in the rare events logistic regression analyses that has a cultural character and is not directly inherent to the physical environment. Digital records of other socio-economic variables (e.g. secondary roads network, political borders, sites with a religious nature) are either unavailable or spatially incomplete and were therefore not included in the analyses. The spatial distribution of contemporary agricultural land use and potential soil erosion is assumed to affect the discovery of archaeological finds regardless of their age, reflected by prospection and preservation biases. These two variables were hence tested as an explanatory variable per region and site type, integrated over the three time periods. As the effective cumulative erosion since the Roman Period is the product of potential soil erosion and an unknown history of land use, this quantity could not be incorporated in the regression analyses.

Archaeological site type	Early Roman Period (ca. 50 BC - AD 270)	Late Roman Period (ca. AD 270 - AD 450)	Merovingian Period (ca. AD 450 - AD 751)
Hesbaye (3630 km ²)			
Burial sites	202	(10)	5 ⁰
Habitation sites	168	(25)	(13)
Condroz (2720 km ²)			
Burial sites	57	(7)	5 ⁸
Habitation sites	54	(17)	(8)

Table 1. Number of observations per combination of archaeological site type, period and region. Combinations with less than 50 observations (between brackets) did not serve as a dependent variable in the rare events logistic regression analyses.

Sample points to calibrate the rare events logistic regression models were, as recommended by King and Zeng (2001b), selected according to a stratified sampling scheme in which all positive observations of the rare-event variable and a random selection of non-events or control points are included. Although research in ecological modelling (Phillips *et al.* 2009; Hastie & Fithian 2013) has indicated that true-absence models are more robust than presence-only approaches to sample the dependent variable, the lack of certitude about the actual absence of an archaeological or occupation site does indeed necessarily involve a random sampling scheme, thus increasing chances of inclusion of false negative observations. The control sample constitutes

Explanatory variable	Description
Elevation (m a.s.l.)	Derived from SRTM-DEM (3"; © NASA and NGA, 2008); bilinearly resampled to 100 m spatial resolution
Slope (%)	Derived from SRTM-DEM
Aspect (°)	Derived from SRTM-DEM
Aspect class	Derived from SRTM-DEM; <i>Classes:</i> NW-NE, NE-SE, SE-SW, SW-NW
Local relief (m)	Derived from SRTM-DEM; defined as the on-site elevation subtracted by the average elevation within a 2-km radius
Absolute value of local relief (m)	Derived from SRTM-DEM
Landform class	Derived from SRTM-DEM ; classified according to Morgan and Lesh (2005), based on Dikau's (1989) algorithm; <i>Classes:</i> plains, tablelands, hilly plains, open hills, hills
Incoming solar radiation (WH m ⁻²)	Derived from SRTM-DEM; calculated on a yearly basis using Esri ArcGIS*
Lithology class	Derived from the vectorised geological maps of Belgium (1:40,000; © Geological Survey of Belgium, 2007); defined as the dominant lithology class within a 500-m radius; cf. Fig. 2; <i>Classes:</i> alluvial and colluvial deposits, sand, clay, claystone (including schist), sandstone (including psammite and quartzite), limestone (including chalk, tufa and marble), mixed lithology, other
Distance to lithology classes (m)	Derived from the vectorised geological maps of Belgium
Distance to border between pairs of neighbouring lithology classes (m)	Derived from the vectorised geological maps of Belgium
Soil drainage class	Derived from the digital soil maps of Flanders (1:20,000; © AGIV, 2001) and Wallonia (1:20,000; © SPW, 2009); defined as the dominant drainage class within a 500-m radius; <i>Classes:</i> dry, moderately humid, humid to very humid
Soil texture class	Derived from the digital soil maps of Flanders and Wallonia; defined as the dominant texture class within a 500-m radius; <i>Classes:</i> loam, sandy loam, sand, clay, stony
Soil suitability for wheat and for pasture	Derived from the digital soil maps of Flanders and Wallonia in combination with suitability values ranging between 1 (highly suitable) and 5 (unsuitable), which are reported per soil type in explanatory notes accompanying the Belgian soil maps; defined as the average suitability within a 500-m radius
Distance to rivers (m)	Derived from SRTM-DEM; calculated on a river network generated through a runoff algorithm with an upstream contributing area threshold of 50 ha (Verstraeten, 2006)
Distance to Roman roads (m)	Derived after digitalisation of major Roman roads that are attested by excavations or are generally believed to have existed based on historical maps, written sources or toponyms (e.g., Mertens, 1957; Corbiau, 2005; Corbiau, & Yante, 2010); cf. Fig. 1
Presence of arable land or pasture	Derived from the raster-based land cover map of Flanders (15 m; © AGIV, 2002) and the vector-based cadastral land cover data of Wallonia (© SPW, 2007)
Potential erosion E (t ha ⁻¹ yr ⁻¹)	Calculated based on the Revised Universal Soil Loss Equation $E = R K L S C$ (Wischmeier, & Smith, 1978); $R = 877 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ yr}^{-1}$ (Verstraeten et al., 2003); K was obtained from the European Soil Database (500 m; © European Commission, 2015); LS (m m ³) was calculated based on the SRTM-DEM in WATEM/SEDEM (Van Oost, Govers, & Desmet, 2000; Van Rompaey et al., 2001; Verstraeten et al., 2002); $C = 1$

Table 2. Explanatory variables tested in the logistic regression analyses.

approximately 4% of the total area to be sampled, corresponding to 10,000 and 14,000 points for the Condroz and the Hesbaye, respectively. The random-sample density within each region was varied depending on the quality or completeness of the archaeological survey on the municipality level, where well-prospected areas are sampled more densely in order to minimise the erroneous selection of yet unknown sites as non-events. For the same reason, taking into account both the geographical inaccuracy of reported site locations and the spatial extent of sites, a 500-m radius circular buffer zone around each site was excluded from the area to be sampled.

Prior the regression analyses, a Spearman correlation test was applied to the calibration sets. Explanatory variable pairs with a correlation coefficient exceeding 0.5 were flagged as highly correlated and are in further analyses inhibited to act together as an independent variable in a regression model.

Univariate rare events logistic regression analyses were performed for each explanatory variable in order to assess the influence of factors that might not be included in the best-performing multivariate model. Next, bidirectional stepwise multivariate *ordinary* logistic regression models were fitted to find the optimal combination of explanatory variables to describe each dependent variable. Inclusion or exclusion of the tested variables in the stepwise selection procedure is based on minimisation of Akaike's (1973) Information Criterion (AIC), a relative measure of model performance that takes into account both the model likelihood and the number of parameters. Factor combinations performing well for a given dependent variable were subsequently tested for the same response variable in a rare events regression analysis and evaluated based on the explanatory variables' z-scores and the criterion of absence of correlated variables. The upper threshold of the probability that the parameter estimate statistically equals zero and the model is hence not significant, was set at 0.05 in all regression models. Final model selection was based on maximisation of the area under the receiver operating characteristic (ROC) curve, which plots for different probability cut-off values the proportion of correctly predicted events relative to the total number of positive observations in the sample set, against the proportion of wrongfully predicted events relative to the total number of non-events (Fawcett 2006). ROC-values exceeding 0.5 indicate that the model performs better than a random model. Regression parameter estimates and the corresponding odds ratios give an indication on the direction and magnitude of the variables' influence on site occurrence. The odds ratio of an explanatory variable with parameter estimate $\hat{\beta}$ is calculated as $e^{\hat{\beta}}$ and is a measure of change in the probability $P(Y = 1)$ with each unit increase of x , when all other factors are constant. Hence, odds ratios above 1 indicate a positive relation, whereas values between 0 and 1 point to a negative influence of the variable on the estimated site occurrence probability. If x is a dummy variable, the odds ratio expresses how the probability for observations where $x = 1$ is related to the probability if $x = 0$.

Finally, the selected models' parameter estimates served, together with the explanatory variable images, to create maps depicting site occurrence probabilities for the entire Condroz and Hesbaye based on Eq. 1. Frequency density curves of site probability values in the entire region, at site locations only, and the average within a 500-m radius around sites, were compared in order to evaluate the constructed probability maps. Analogous to the selection criterion of the control sample, this site radius was introduced to account for locational inaccuracies and the sites' spatial dimensions.

RESULTS

Rare events logistic regression analyses

Results of the univariate rare events logistic regression analyses are summarised in Table 3. The statistical significance of multiple variables that reflect the landscape's topography, the proximity of water, and soil fertility, indicates that these physical properties of the environment are important factors with respect to the location of burial and habitation sites, although their influence may vary with different sets of sites. The proximity of the main Roman roads and the presence or proximity of a limestone substrate are significantly positively related to archaeological sites in all studied periods and regions.

Tables 4 and 5 present, respectively for the Hesbaye and the Condroz, the explanatory variables of the best-performing multivariate rare events logistic regression model for each combination of site type and period, together with the corresponding parameter estimates and significance and performance measures. According to the ROC-values, which all exceed the 0.5 threshold associated with a random model, the logistic relation that describes the occurrence of Merovingian burial sites in the Condroz is by far the best-performing model, followed by the Hesbaye model for this site type and period, despite the relatively small number of positive observations. The regression models for Early Roman burial and habitation sites in the Hesbaye perform almost equally well, whereas the occurrence of Roman sites in the Condroz is to a lesser degree determined by the tested explanatory variables. The selected variables in each model correspond largely with the outcomes of the univariate regression analyses (tbl. 3). Yet, in some cases, the interaction between independent variables may lead to the selection of a new, perhaps counterintuitive set of variables (see, e.g., the pronounced positive effect of clay, sand and alluvium in Merovingian burial site probability calculations). Explanatory factors in a multivariate regression model should hence always be interpreted in relation to one another.

Explanatory variable	Hesbaye			Condroz		
	E.R. burial	Merov. burial	E.R. habit.	E.R. burial	Merov. burial	E.R. habit.
Elevation	+	NS	NS	-	-	NS
Slope	NS	NS	+	NS	+	NS
Aspect	+	NS	NS	NS	NS	NS
Aspect class						
NW-NE	NS	NS	NS	NS	NS	NS
NE-SE	-	NS	NS	NS	NS	NS
SE-SW	NS	+	NS	NS	NS	NS
SW-NW	+	NS	NS	NS	NS	NS
Local relief	+	-	NS	NS	-	NS
Absolute value of local relief	NS	+	NS	NS	+	-
Landform class						
Plains	+	+	+	NS	NS	NS
Tablelands	+	+	NS	NS	NS	NS
Hilly plains	NS	+	+	NS	+	NS
Open hills	NS	NS	NS	NS	NS	NS
Hills	NS	NS	NS	NS	NS	NS
Incoming solar radiation	+	NS	NS	-	NS	NS
Lithology class						
Alluvium	-	NS	NS	+	+	NS
Sand	+	-	NS	/	/	/
Clay	NS	NS	NS	/	/	/
Claystone	+	+	+	NS	NS	NS
Sandstone	NS	NS	NS	NS	-	NS
Limestone	NS	NS	+	+	+	+
Mixed lithology	-	NS	-	NS	-	-
Distance to alluvium	+	NS	+	NS	-	NS
Distance to limestone	-	-	-	NS	-	NS
Distance to sandstone	/	/	/	NS	+	NS
Distance to alluvium-sand border	NS	+	+	/	/	/
Distance to alluvium-sand and alluvium-sandstone border	NS	+	+	/	/	/
Distance to alluvium-clay border	NS	+	NS	/	/	/
Distance to alluvium-claystone border	/	/	/	NS	NS	NS
Distance to alluvium-sandstone border	+	NS	+	NS	NS	NS
Distance to alluvium-sandstone and alluvium-claystone border	/	/	/	NS	-	NS
Distance to alluvium-limestone border	-	-	-	-	-	NS
Distance to sand-clay border	NS	+	NS	/	/	/
Distance to limestone-sand border	-	-	-	/	/	/
Distance to limestone-sand and limestone-sandstone border	-	-	-	/	/	/
Distance to limestone-claystone border	/	/	/	-	NS	NS
Distance to limestone-sandstone border	/	/	/	NS	-	NS
Distance to limestone-sandstone and limestone-claystone border	/	/	/	-	-	NS
Distance to sandstone-claystone border	/	/	/	NS	NS	NS
Soil drainage class						
Dry	+	+	+	NS	NS	NS
Moderately humid	-	NS	-	NS	NS	NS
Humid to very humid	-	NS	-	+	+	+
Soil texture class						
Loam	+	NS	+	NS	-	NS
Sandy loam	-	NS	-	/	/	/
Sand	NS	NS	NS	/	/	/
Clay	+	NS	+	+	+	+
Stony	NS	+	+	NS	+	NS
Soil suitability for wheat	-	-	-	NS	NS	NS
Soil suitability for pasture	-	NS	-	NS	+	NS
Distance to rivers	+	NS	NS	NS	-	NS
Distance to Roman roads	-	-	-	-	-	-

Table 3. Univariate rare events logistic regression model results for Early Roman (E.R.) burial and habitation (habit.) sites and Merovingian (Merov.) burial sites in the Hesbaye and Condroz regions. (+ : positive relation; - : negative relation; NS: not significant at a 0.05 level; /: the variable does not occur in the region or in the sample. Note: the variables expressing soil suitability for wheat and for pasture are scaled reversely from 5 to 1. Their influence on site occurrence should be interpreted accordingly.)

Model	Parameter estimate	Standard error	z-value	Pr(> z)	Odds ratio
Early Roman burials (n = 202, ROC = 0.716)					
Intercept	-6.545E+00	2.185E-01	-2.996E+01	< 2.00E-16	
Aspect SW-NW	6.582E-01	1.471E-01	4.474E+00	7.690E-06	1.93
Local relief	4.096E-02	8.119E-03	5.045E+00	4.530E-07	1.04
Soil suitability for wheat	-6.247E-01	1.255E-01	-4.978E+00	6.410E-07	0.54
Distance to alluvium-sandstone border	5.847E-05	1.258E-05	4.646E+00	3.380E-06	1.00006
Distance to alluvium-sand border	-1.360E-04	3.195E-05	-4.255E+00	2.090E-05	0.99986
Distance to limestone-sand border	-3.942E-05	1.288E-05	-3.061E+00	2.210E-03	0.99996
Merovingian burials (n = 50, ROC = 0.725)					
Intercept	-1.066E+01	1.052E+00	-1.013E+01	< 2.00E-16	
Aspect SE-SW	7.433E-01	3.054E-01	2.434E+00	1.492E-02	2.10
Soil drainage dry	2.183E+00	1.031E+00	2.118E+00	3.421E-02	8.87
Soil texture sand	1.627E+00	7.782E-01	2.090E+00	3.659E-02	5.09
Soil texture clay / peat	3.107E+00	1.445E+00	2.151E+00	3.149E-02	22.35
Alluvium	1.585E+00	4.846E-01	3.272E+00	1.070E-03	4.88
Distance to limestone-sand and limestone-sandstone border	-9.627E-05	3.107E-05	-3.098E+00	1.950E-03	0.99990
Early Roman habitation (N = 168, ROC = 0.721)					
Intercept	-9.298E+00	4.476E-01	-2.078E+01	< 2.00E-16	
Slope	7.443E-02	2.025E-02	3.675E+00	2.370E-04	1.08
Soil texture loam	1.337E+00	3.931E-01	3.400E+00	6.730E-04	3.81
Mixed lithology	-1.740E+00	5.117E-01	-3.401E+00	6.720E-04	0.18
Distance to alluvium-sandstone border	6.946E-05	1.095E-05	6.346E+00	2.210E-10	1.00007
Distance to limestone-sand border	-4.518E-05	1.407E-05	-3.211E+00	1.325E-03	0.99995

Table 4. Multivariate rare events logistic regression model results for three combinations of archaeological site type and period in the Hesbaye region (n: number of observed sites; Note: the variable expressing soil suitability for wheat is scaled reversely from 5 to 1. Its influence on site occurrences should be interpreted accordingly).

Model	Parameter estimate	Standard error	z-value	Pr(> z)	Odds ratio
Early Roman burials (n = 57, ROC = 0.647)					
Intercept	-8.044E+00	2.528E-01	-3.182E+01	< 2.00E-16	-
Aspect SE-SW	-7.265E-01	3.643E-01	-1.994E+00	4.616E-02	0.48
Distance to alluvium-limestone border	-5.164E-04	1.466E-04	-3.522E+00	4.280E-04	0.99948
Distance to alluvium-sandstone border	2.287E-04	7.886E-05	2.900E+00	3.735E-03	1.00023
Merovingian burials (n = 58, ROC = 0.784)					
Intercept	-8.026E+00	3.640E-01	-2.205E+01	< 2.00E-16	-
Local relief	-1.474E-02	4.912E-03	-3.001E+00	2.690E-03	0.99
Landform hilly plains	9.234E-01	3.680E-01	2.510E+00	1.208E-02	2.52
Soil texture stony	6.702E-01	3.388E-01	1.978E+00	4.793E-02	1.95
Distance to alluvium-limestone border	-9.669E-04	2.086E-04	-4.635E+00	3.570E-06	0.99903
Early Roman habitation (n = 54, ROC = 0.671)					
Intercept	-7.963E+00	3.187E-01	-2.499E+01	< 2.00E-16	-
Absolute local relief	-2.377E-02	1.237E-02	-1.922E+00	5.459E-02	0.98
Distance to alluvium-limestone border	-4.129E-04	1.529E-04	-2.700E+00	6.940E-03	0.99959
Distance to alluvium-sandstone and alluvium-claystone border	3.074E-04	1.441E-04	2.133E+00	3.296E-02	1.00031

Table 5. Multivariate rare events logistic regression model results for three combinations of archaeological site type and period in the Condroz region (n: number of observed sites).

Figures 4 and 5 depict the modelled site occurrence probability patterns, which reflect the combined effect of the selected explanatory variables, as well as the frequency distribution of probability values in the region, at site locations, and in a 500-m radius around each site. In absolute terms, the modelled probabilities of site occurrence range from approximately zero to a maximum of 0.007 in the Condroz and 0.045 in the Hesbaye. As their magnitude strongly depends on the total number of positive observations in the study area, such low probability values are common for rare events data, even after the applied corrections. However, when employed in an archaeological context where events and non-events are largely determined by human decisions that cannot be controlled or predicted and where site data are moreover incomplete, the modelled probability

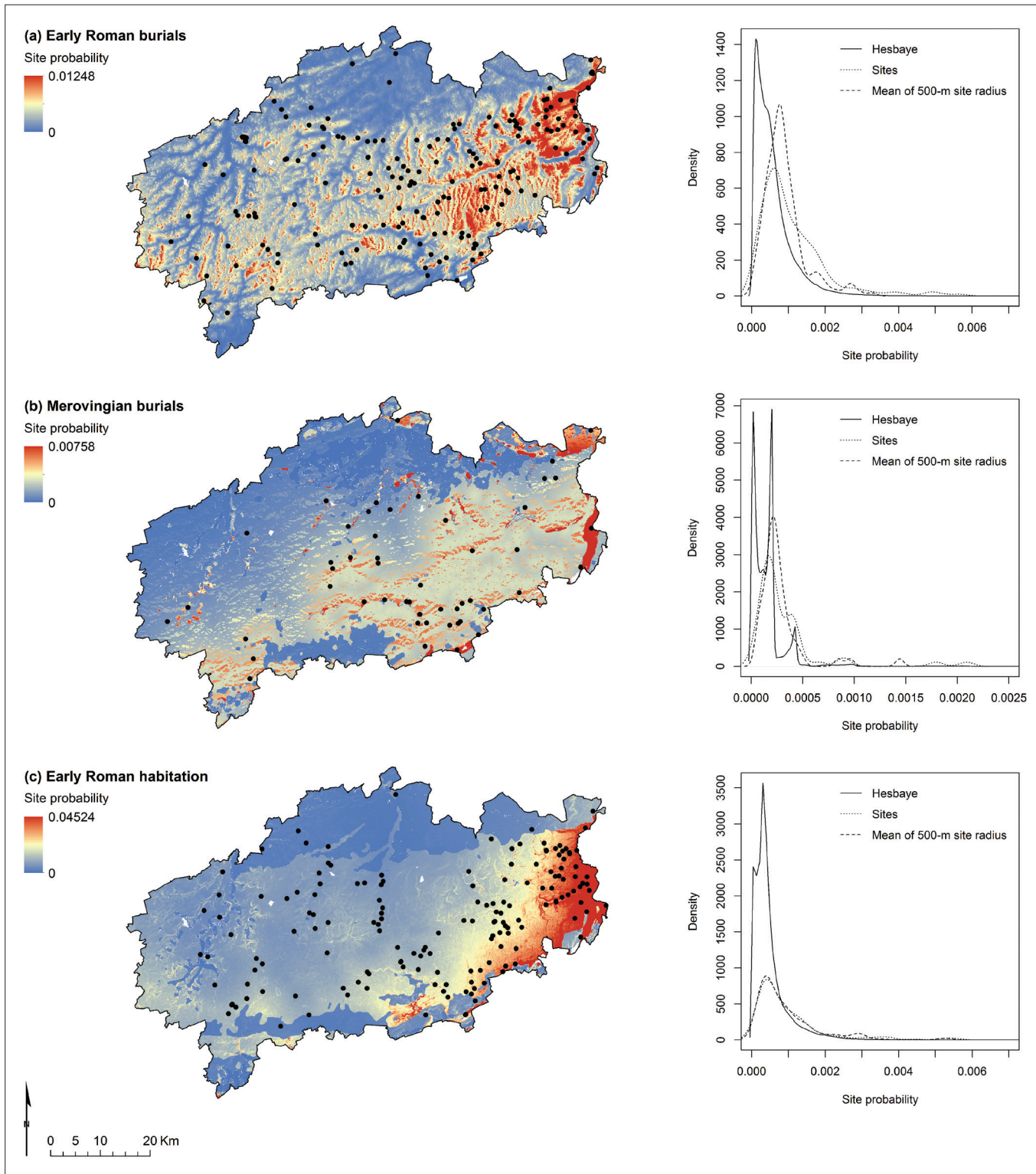


Fig. 4. Site occurrence probability maps (left) and frequency density plots (right) for three combinations of archaeological site type and period in the Hesbaye region, based on multivariate rare events logistic regression models. Black dots represent recorded site locations in each model. Frequency density distributions of the modelled probabilities were plotted for the entire region, at site locations, and for a 500-m radius around sites (mean value of the radius).

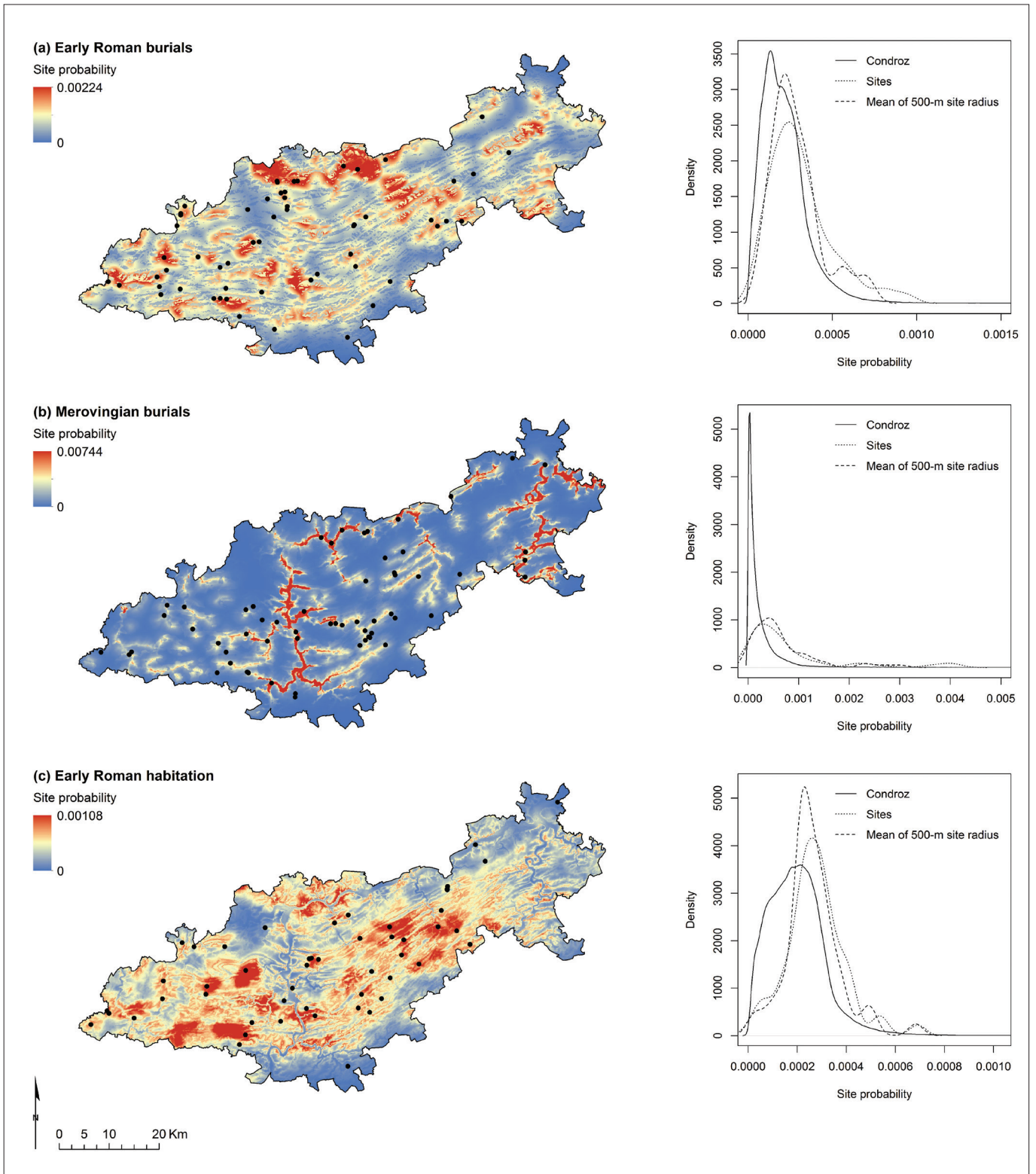


Fig. 5. Site occurrence probability maps (left) and frequency density plots (right) for three combinations of archaeological site type and period in the Condroz region, based on multivariate rare events logistic regression models. Black dots represent recorded site locations in each model. Frequency density distributions of the modelled probabilities were plotted for the entire region, at site locations, and for a 500-m radius around sites (mean value of the radius).

values and their spatial distribution are to be interpreted relative to each other rather than in an absolute sense. Visual analysis of observed site patterns of a given type and period plotted onto their respective occurrence probability maps suggests a good agreement between reported site locations and areas with relatively high probabilities, in particular for Merovingian burial sites in the Condroz, but also for settlements in this region and for Early Roman burial sites in the Hesbaye. The correspondence is less clear for Merovingian burial and Early Roman habitation sites in the Hesbaye and for Early Roman burial sites in the Condroz, despite good ROC-values of the former two. Marked low-probability zones show little to no positive observations in all cases.

Frequency density plots of the probability values confirm the good performance of the logistic regression model for Merovingian burial sites in the Condroz, and to a lesser extent of the Early Roman Condrosian habitation site and Hesbayian burial site occurrence models. The probability mode of sites and site radii indeed occurs at significantly higher values than the regional peak, although there is still an important overlap. In contrast, the density curves of settlements and Merovingian burial sites in the Hesbaye and Early Roman burials in the Condroz show less distinct site probability distributions. Multimodality in the frequency distributions, particularly evident in the Merovingian burial site probabilities for the entire Hesbaye, can be related to the discrete nature of categorical factors that determine the logistic relation. Numerical explanatory variables, in contrast, yield smoother curves. Note that the discrete or continuous character of the variables' value range is also reflected in the spatial pattern of site occurrence probabilities. The probability frequency densities of the site locations and those of the mean value within a larger area around each site are highly similar in most models, and one does not systematically outperform the other. The relation between both depends on the spatial variability of the explanatory variables at the site radius scale, as well as on the spatial accuracy of the site locations.

Modern land use and potential soil erosion in relation to site density and probability

The impact of agricultural land use and potential soil erosion on the modelled occurrence probability and observed site density was first investigated by assessing their significance, along with all other explanatory variables listed in Table 2, in multivariate regression analyses. All recorded sites dating from the Early Roman up to the Merovingian Period served to compile four dependent variables, including 262 burial sites and 206 settlements for the Hesbaye, and 122 burial sites and 79 habitation sites for the Condroz. ROC-values ranging between 0.712 and 0.731 indicate that the resulting multivariate models perform well. With respect to the selected explanatory variables, the obtained logistic relations are for both regions highly similar to their counterparts predicting site occurrences for only one period (tbl. 4 and 5), and are therefore not presented here in further detail. The presence of arable land and pastures appears to negatively influence the probability of burial sites in the multivariate models for both regions. In addition, univariate regression analyses indicate that the occurrence of settlements is positively related to the presence of agricultural land in the Condroz, and to potential soil erosion in the Hesbaye. The latter variable did not show a significant effect in any other tested model. Note that the significance and direction of the impact of agricultural land on site occurrences is directly linked to the number of sites per unit area for the entire region and for the area under agriculture.

The histograms in Figure 6 present the area, mean modelled site occurrence probability and observed site density for different classes of potential soil erosion in the Hesbaye and Condroz, with a differentiation between regional values and those calculated for the area under agriculture only. The areal distribution of potential erosion classes (fig. 6a, f) indicates that the share of potentially more erodible soils is small in both regions. The distribution of arable land and pastures does not differ notably from the regional image. Modelled burial site probabilities in the Hesbaye (fig. 6b) vary only slightly with potential soil erosion, whereas probabilities in the Condroz (fig. 6g) increase markedly with the soil's erosion sensitivity. In both regions, site probabilities at locations with low potential erosion are on average larger on agricultural land compared to values for the entire region, while the opposite is true for the largest erodibility classes on steep valley slopes. Unlike their probability distribution, observed densities of burial sites in the Hesbaye (fig. 6c) show a sharp increase with potential soil erosion. About burial site densities, the relation between regional and agricultural area for different potential erosion classes contrasts with the observed trend for modelled probabilities in the Hesbaye. The positive effect of the soil's erodibility on the density of Condrosian burial sites is not as pronounced (fig. 6h).

Modelled habitation site probabilities in the Hesbaye (fig. 6d) are highly positively – but not necessarily causally – related to potential soil erosion. This is also the case for agricultural land, albeit to a lesser degree. Apart from a local minimum in the observed site density at potential erosion values between 30 and 40 t ha⁻¹ yr⁻¹, the relation between Hesbayian settlement densities and erosion sensitivity (fig. 6e) corresponds well with the modelled probability distribution of Figure 6d. In the Condroz, both regional and agricultural settlement probabilities are quasi-invariable with potential erosion (fig. 6i), whereas the corresponding site densities show local maximums at the intermediate as well as the most extreme erosion sensitivity classes (fig. 6j). Remarkably, both settlements and burials occur relatively frequently at potential erosion values between 30 and 50 t ha⁻¹

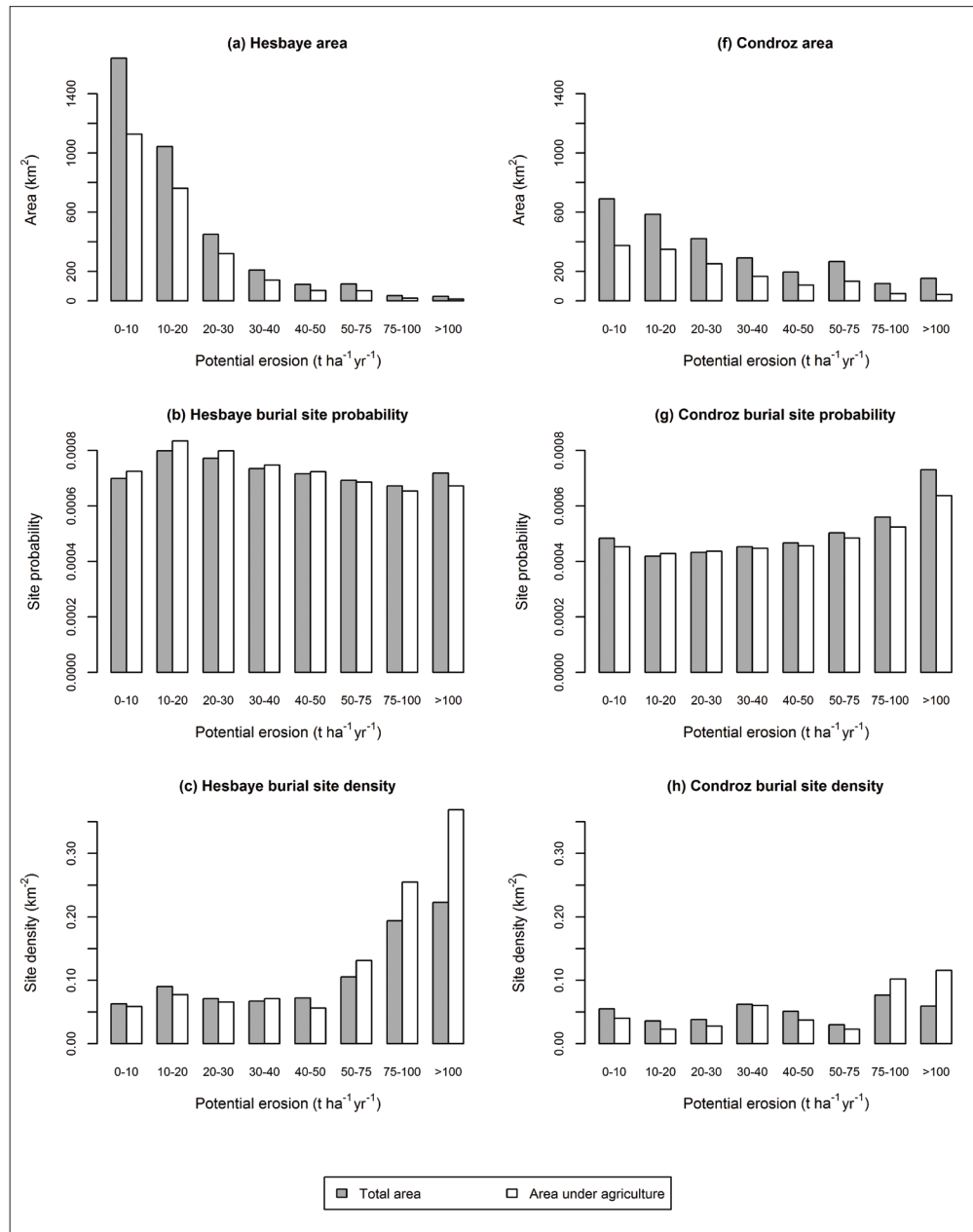


Fig. 6. Area, mean modelled site occurrence probability, and observed site density per potential erosion class for the Hesbaye (left) and Condroz (right) regions. All recorded burial and habitation sites from the Early Roman to Merovingian Period were taken into account. Histograms were created for the entire region (gray bars), as well as for cultivated areas only (white bars).

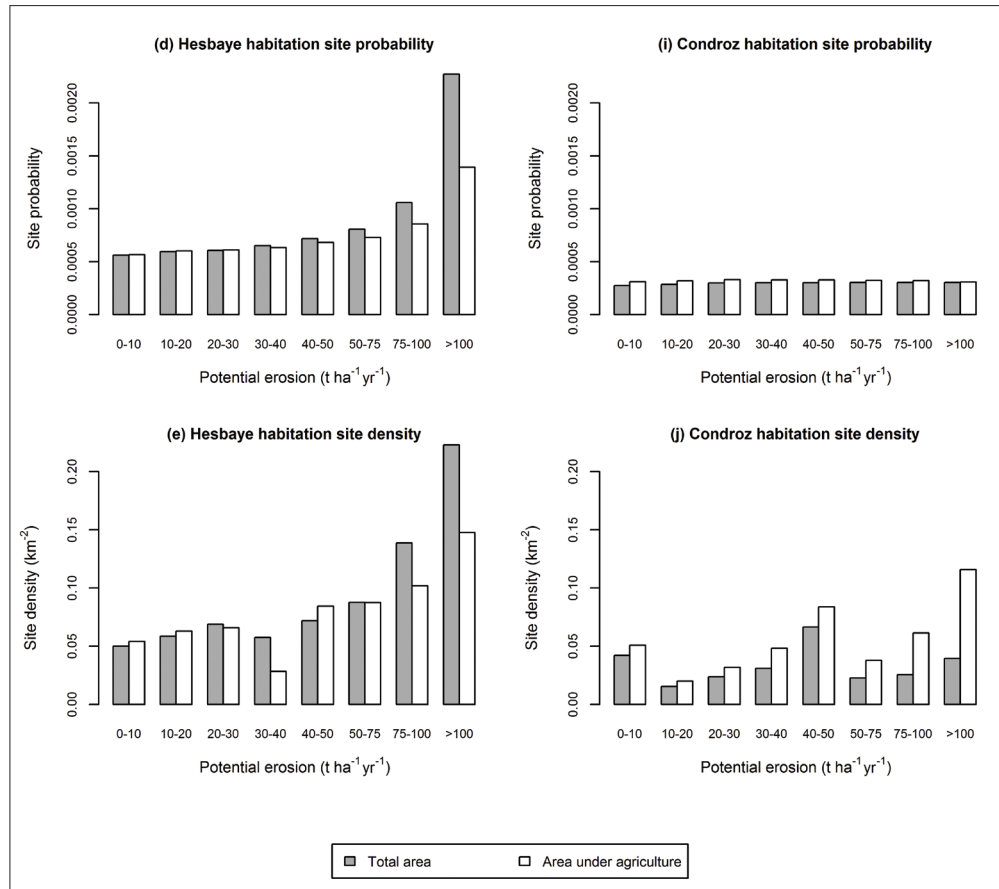


Fig. 6. (continued)

yr⁻¹ in the Condroz. This trend was not captured in the regression models and may reflect a positive preservation or discovery bias in a specific physical setting that is characterised by intermediate potential erosion rates. Settlement densities under agricultural land in this region exceed the values for the entire Condroz, and the difference between both becomes larger as the erosion potential increases.

DISCUSSION

Model performance

Given the incompleteness of archaeological site data and the rare-event nature of the data set, an evaluation of the constructed probability models is not straightforward. Moreover, frequently applied logistic-model evaluation and validation measures such as Cohen's (1960) kappa index and the total agreement proportion, are applicable to binary variables only and would hence require conversion of probability values to site and non-site classes, as well as the availability of a validation sample or ground truth image. Nevertheless, a threefold assessment of the multivariate regression models' performance based on their ROC-value (cf. tabl. 4 and 5), the modelled probability patterns and the frequency density curves of regional, on-site and near-site probabilities (fig. 4 and 5), indicates that the location of some of the tested sets of sites is indeed to a large extent determined by the physical environment. However, several models' ROC-values conflict with the density curves and the comparison of observed site locations with mapped probabilities. This is the case for Early Roman habitation and Merovingian burial sites in the Hesbaye, which show surprisingly large ROC-values. For instance, the regional probability pattern for Early Roman Hesbayian settlements (fig. 4c) appears to overemphasise the eastern part of the region, thereby misleadingly suggesting that

sites are unlikely to occur in the central and western Hesbaye. On the other hand, the Early Roman Condrusian settlement model is characterised by a moderate ROC-value despite its otherwise positive evaluation (fig. 5c).

A possible explanation for the perceived inconsistency between evaluation methods is the sensitivity of the ROC to the proportion of observed sites relative to the *true* number of sites, both discovered and undiscovered. In the archaeological record, negative “observations” are in fact determined by the lack of a positive observation – caused by poor prospection or preservation – and therefore highly likely to contain false negatives. Consequently, the calculated area under the ROC-curve is necessarily an underestimation, as the actual proportion of true positives and false positives shows respectively an increasing and decreasing relation with the number of undiscovered sites. In contrast to Early Roman Hesbayian sites, positive observations are indeed scarce in the Condruz, resulting in a relative underestimation of the latter model’s ROC-based performance. The small sample size of the Merovingian burial site model in the Hesbaye, however, calls for another explanation and suggests that its seemingly large ROC-value is partly based on chance or overfitting of the calibration data, rather than on a true causal connection between site locations and the selected combination of explanatory variables (cf. tabl. 4). In particular, due to the statistical unlikelihood of archaeological sites to occur on a spatially underrepresented area, a sole positive observation on a scarce clayey patch or in the limited sandy area in the northern part of the Hesbaye suffices to yield peculiarly high probability values in these zones (see fig. 4b). It is, however, not likely that the chance of discovering Merovingian burial sites is indeed much larger in these areas as compared to loess-covered soils. A final remark concerning the ROC-value is that this measure, in contrast to the density curves and probability maps, is based on calibration data only, which cover *ca.* 4% of the study area and were sampled with varying densities depending on the archaeological prospection quality. As such, it might deviate significantly from model evaluations based on the true population’s probability distribution.

On the other hand, evaluation based on density plots of the modelled probabilities builds on the idea that well-performing models have their on-site and near-site frequency modes at significantly larger probability values compared to the regional distribution maximum, whereas considerable overlap of the curves is perceived as undesirable. Although such overlap can indeed be a sign of limited model performance (i.e., the selected set of explanatory variables does not cover all factors that determine site location), it can also simply indicate that the area suitable for occupation is much larger than the area that is actually covered by (observed) sites. Hence, the smaller the portion of the study area constituted by high site occurrence probabilities, the less overlap the density curves will show, provided that the logistic model successfully captures the relation with the explanatory variables.

Finally, evaluation based on a visual comparison of observed site locations and modelled probability patterns is by definition subjective and might even be misleading. Areas characterised by notably low probabilities are indeed quasi-void of observations in all tested models, though, hence contributing to a positive evaluation.

Functional, temporal and regional comparison of modelled site probabilities

Despite varying degrees of performance and issues related to model evaluation, the tested logistic regression models do allow to assess several functional, temporal and regional differences and similarities in site occurrences and their determining factors, discussed in the following sections. For this purpose, the univariate models have proven to provide a useful addition to the multivariate models, facilitating their interpretation and allowing to evaluate the isolated influence of various explanatory variables.

Burial vs. habitation sites

The preferential allocation patterns of burial and habitation sites in the Hesbayian and Condrusian landscapes encompass some generalities. Burial sites tend to be located in the valleys of either small or larger rivers, at relatively low altitude and in relief-rich, topographically rather closed areas. Merovingian sites, specifically, are linked to the presence of stony soil textures. Burial sites of this period show the strongest relation with the physical environment of all tested models, despite their limited number of observations. Monumental graves built according to the (Early) Roman tradition reflect a criterion of visibility (e.g. Ferdière 2004; Crowley 2011; see also: Massart 1994; Massart 2015), and hence deviate from the general burial site pattern. The large proportion of tumuli (*ca.* 63%) in the Early Roman burial site record of the Hesbaye indeed explains the tendency towards allocation on the more elevated plateaus in this model. A separate set of multivariate regression analyses,

i.e., one for all Early Roman burial sites in the Hesbaye except tumuli ($n = 71$, $ROC = 0.667$) and one for tumuli only ($n = 131$, $ROC = 0.783$), confirms this assumption. In the Condruz, tumuli constitute only *ca.* 17% of the observed Early Roman burial sites. Their apparent effect on estimated site probabilities illustrates the importance of careful definition of the predictor variable (Kohler & Parker 1986): if the latter combines features that reflect two or more sets of localisation factors, these cannot be accurately captured in one logistic model. Similarly, the application of a single logistic relation to fit all recorded rural settlements obscures any differentiation in, e.g., the social status, political and religious beliefs and main economic activities of their inhabitants. Information on such discrepancies in allocation processes is however not always a priori available, and could potentially bias model outcomes.

Rural settlements in our study area, investigated for the Early Roman Period only, are generally situated at higher elevations and in more open landscapes compared to burial sites. Unlike Condruzian settlements, habitation sites in the Hesbaye may also occur on the valley slopes of the Dijle and Gete rivers and some of the larger tributaries of the Meuse (e.g. Jeker, Méhaigne). The quasi-absence of settlements on major interfluves further illustrates the importance of water availability with respect to site localisation. The potential concurrence of administrative boundaries with these watersheds (as well as other natural landmarks notably) may play a role as well (Cardot 1987; Raepsaet-Charlier 1994; Noël 1991; Verslype 2003b) in certain cases, but not exclusively (Guizard-Duchamp 2003).

Note that the limited number of Merovingian habitation sites in the archaeological record (i.e., only 20 sites for the entire study area) can at least partially be attributed to the poor preservation potential of wooden Germanic farms, as the number of burial sites does suggest that the Hesbaye and Condruz were relatively densely populated during this period. Thus, even though both types of sites may be characterised by different location preferences at local scales (e.g. at the hillslope or pixel level), burial sites are a valuable proxy for human occupation patterns at the subregional scale.

The presence and proximity of limestone has proven to be a key factor in modelling occurrence probabilities of both burial and habitation sites in the study area. In the Hesbaye, site patterns are in addition highly determined by the spatial extent of the Loess Belt. The fact that limestone is in all tested models positively related to the presence of Hesbayian sites is remarkable, as the substrate lithology is not expected to significantly affect soil fertility in a region covered with thick loess deposits, in contrast to the Condruzian landscape. It is indeed more likely that the perceived relation with the chalky substrate in the former region is in fact a reflection of the proximity of the *caput civitates*, i.e., Tongeren, which is situated near the northern border between limestone and Tertiary sands (cf. fig. 2) and is not directly represented in the list of explanatory variables. Figure 1 confirms that the density of the archaeological site record is disproportionately high in the vicinity of Tongeren.

Literature on the Roman and Merovingian occupation in the study area supports several relations described by the regression models, including the preference for loess soils in the allocation of Early Roman villas that contrasts with the more northerly situated sandy-loamy and sandy areas and matches the general perception of agriculture as a key element in Roman culture and economy (e.g. Nouwen 2006). Also related to the agriculture-oriented economy of Roman and Merovingian times, is the allocation of settlements near the market and major roads or waterways (Nouwen 2006; Roymans & Derks 2011). The univariate regression analyses indeed indicate a significant relation with the distance to Roman roads for each tested combination of site type, period and region (tabl. 3). The geopolitical and economical choices and the natural variables have simultaneously influenced the land occupation. The identification of the social systems and the market economies in which the settlements are located depends on the analysis of the data collected within each site (pollen, macrorestes, ceramics, fauna). As example, two villas located in Hesbaye, *i.e.* in the same agricultural landscape, can be involved in different local exchange networks with differentiation between the targeted markets (see Pigière & Lepot 2013). Finally, we note that many of the observed relations specific to Merovingian burial sites correspond well with Roosens' (1949) conclusions following a qualitative study of these sites for Belgium, specifically the preferential localisation on all but stony soils near headwaters or main river reaches, in proximity to land suitable for agriculture.

Early Roman (vs. Late Roman) vs. Merovingian Period

The decline and decentralisation of economic and social structures that presumably characterised the transition from Roman to Merovingian times in northern Gaul caused a shift in agricultural and other commercial activities, the role of urban centres, transport routes and political borders herein, as well as in the perception of cultivated space and the environment in general (e.g. Vanpoucke *et al.* 2007). Such cultural changes likely also affected the spatial behaviour of Roman and Merovingian societies in our study area. Yet, the extent to which these discontinuities are reflected in the archaeological footprint, remains hitherto unclear.

A diachronic study of the factors influencing site patterns based on the presented logistic models is only possible for Early Roman to Merovingian burial sites. Moreover, the small number of Hesbayian Merovingian-aged burial sites further complicates an analysis of the temporal evolution of site locations, and so does the bias associated with the large share of tumuli in this region. Hence unsurprisingly, similarities in localisation factors between the two analysed periods are particularly evident in the Condroz. Environmental characteristics that appear to be important through time include river and limestone proximity, as well as the presence of loess in the Hesbaye. The positive influence of stony soils on Merovingian burial sites in both regions was not detected for Early Roman graves. Nevertheless, despite the observed continuity in (the interpretation of) several explanatory variables, the corresponding probability maps (fig. 4a-b and 5a-b) exhibit considerably distinct patterns, indicating that other independent variables as well as the parameter estimates of common explanatory factors remain highly influential to the spatial distribution of modelled site occurrence probabilities, both at the regional and the local scale.

The archaeological record indicates a temporal continuity of a limited number of burial site locations for both transitions between periods and both regions, whereas the vast majority of the analysed Late Roman settlements were already occupied in the Early Roman Period. Some Early Roman rural settlements, destructed or abandoned in the second half of the 3rd century, were (partly) reoccupied in the 4th, sometimes till the first half of the 5th century. Sometimes, late Roman settlements – documented by pit houses, pits or kilns – moved next to the Roman stone buildings (Van Ossel 1983). Merovingian reoccupations of habitation sites in the record are less numerous or, in the Condroz, not currently known.

In order to include the available records of Merovingian settlements and Late Roman sites in the model-based diachronic analysis of localisation factors, several intermediate rare events logistic regression models that combine site data of two or three consecutive time periods into a single dependent variable, were tested as well. Moderate improvements in model performance and largely unaltered explanatory variable sets (not shown in detail) indicate a persistence in settling choices from Roman to Merovingian times, with a particularly strong similarity within the Roman Period. However, a solid interpretation of such intermediate logistic relations is complicated by the limited number of added site observations with reference to the original probability models of Merovingian burials and Early Roman sites. At the same time, in some cases the yet considerable relative increase in the total number of observations constituting the multi-period dependent variable might in itself cause the perceived increase in model performance (cf. *supra*). Further limitations of the performed analyses stem from the large proportion of tumuli in Early Roman burial sites, particularly in the Hesbaye, and the persistence of sites between periods, yielding duplicate observations that can reinforce the modelled logistic relations for better or for worse.

The Hesbaye vs. the Condroz

The archaeological record of the Hesbaye counts considerably more burial and habitation sites, particularly of Early Roman age, than the Condroz (tabl. 1). This is only partially explained by the larger area covered by the former region: whereas the density of Roman and Merovingian archaeological sites constitutes 0.129 sites km² in the Hesbaye, only 0.074 sites km² were inventoried in the Condroz. The difference in site density may be biased by differential prospection intensities, the current landscape cover, influenced by the vicinity of Tongeren, but can also be related to the particular environmental setting of each region. As noted earlier, Hesbayian sites are characterised by a preferential allocation on loess deposits, while the proximity of limestone is a key factor in the Condroz. Yet, as limestone weathering produces loamy sediments with properties similar to loess, both variables essentially reflect the same location preference. Hence unsurprisingly, archaeological site densities in the Hesbayian loess belt (0.149 km²) and on Condrosian limestone (0.126 km²) do have the same order of magnitude, in contrast to the regional values.

The incorporation of explanatory variables related to the presence of loess and limestone to express the soil's agricultural suitability in the logistic models for the Hesbaye and Condroz, respectively, illustrates how the concrete manifestation of similar site location strategies depends on the physical geography and the natural heterogeneity of these regions. Similarly, the influence of the proximity to rivers in the modelled site probabilities is highly pronounced in the Condroz (e.g. fig. 5b) due to its narrow valleys, which contrast with the relatively wide floodplains of the Hesbayan loess area. A direct consequence of the distinct physical characteristics of both regions is that the occurrence of sites of a specified type and period cannot be captured accurately in a sole statistical model, regardless of potential analogies in the interpretation of the logistic relations. The regression models constructed in this study can for the same reason not be extrapolated to regions with a contrasting physical-geographical structure. This demonstrates that, settlers take advantages to distinct local environmental specificities.

Preservation, prospection and discovery

The representativeness of the archaeological record is a function of the site's preservation and discovery potential as well as prospection biases. Factors determining preservation include not only the spatial imprint and sustainability of the site's structure and materials (e.g. the poor preservation of wooden farms as compared to stone-built villas) and direct anthropogenic impacts, but also geomorphological processes. The rich occupation and land use history in the region has induced significant soil erosion and sediment redistribution since (and already well before) the Roman Period, attested by thick colluvial deposits in lower parts of the landscape and truncated soil profiles and gullies on the valley slopes (Vanwalleghem *et al.* 2006; Rommens *et al.* 2007; Notebaert *et al.* 2011), which might have led to the erosion of material evidence of Roman and Merovingian occupation as well. On the other hand, sites buried under a considerable layer of sediments are generally better preserved but less likely to be discovered. The discovery potential or visibility of sites further depends on modern land use, as well as on the survey techniques and the spatial and thematic focus of archaeological prospection.

In both studied regions, burial sites on agricultural land are relatively less numerous in comparison to uncultivated areas, whereas the opposite is true for Condrosian settlements. However, Figure 6 demonstrated that both observed site densities and the relation between site density and agricultural land use depend on potential soil erosion, and that these relations are not always reflected in the modelled site occurrence probability distributions. A true causal connection between potential erosion and the site density or probability patterns is not guaranteed though. To illustrate, the marked positive relation between modelled Hesbayan settlement probabilities and potential soil erosion (fig. 6d) can either reflect a bias in the sampled sites related to enhanced discovery chances, or it can imply that the modelled (and possibly also the true) spatial allocation preferences are indirectly linked to erosion sensitivity via their explanatory variables (e.g. steep slopes, loamy soils). Nevertheless, the considerable deviation between the modelled probabilities' and the observed site densities' relation with potential soil erosion that is apparent for some combinations of site type and region (compare, e.g., fig. 6.3b and c) despite a positive model evaluation based on their ROC-values, does indicate that the archaeological site record is indeed affected by a spatially variable discovery potential linked with soil erosion.

The connection between soil erosion sensitivity and the preservation or discovery potential of sites is further amplified by the presence of arable land, as this cover type is associated with the highest *effective* erosion rates. Moreover, agricultural land use enhances site visibility also in a direct way, due to the absence of urban or industrial constructions and dense vegetation covers. Hence, the fact that recorded site densities in highly erosion-sensitive areas are, for all sites except Hesbayan settlements, larger on agricultural land than on other land cover types, confirms that the uncovering of sites due to soil erosion increases the chances of discovery. In contrast, the presence of arable land and pastures in the Hesbaye negatively impacts habitation site densities on highly erodible soils compared to the regional average, possibly indicating poor site preservation due to (too) intense effective erosion. For all site types and in both studied regions, the effect of agricultural land on archaeological site occurrence appears less important in areas with low erosion sensitivity. Note that the influence of the type of agricultural land use (i.e., arable land or pasture) on soil erosion and site visibility was not investigated in further detail.

CONCLUDING REMARKS: APPLICATION OF LOGISTIC REGRESSION IN MODELLING ARCHAEOLOGICAL SITE OCCURRENCES

This study investigated the relation between the spatial pattern of Roman to Merovingian burial and habitation sites and various environmental and cultural factors in two contrasting physical environments in Belgium, i.e., the Hesbaye and Condroz regions, through the application of univariate and multivariate rare events logistic regression. Despite varying model performance levels, the explanatory variables significantly influencing archaeological site occurrences and the corresponding site probability maps indicate that burial sites other than tumuli tend to be located in river valleys, whereas settlements are more often situated at relatively elevated, open parts of the landscape. The large-scale site pattern is further affected by the presence or proximity of loess and limestone in the Hesbaye and in the Condroz, respectively. The concrete manifestation of these environmental factors in the logistic regression analyses depends on the physical geography and heterogeneity of the landscape and hence demonstrates the region-specific character of the modelled relations, as well as the aptitude of ancient societies to adapt their spatial behaviour to the local environment. Diachronic analyses of the recorded burial sites and settlements are complicated by the limited number of site observations for some periods and by the bias related to the specific localisation criteria of tumuli in the Early Roman record, but do suggest a temporal continuity of the principal factors determining site patterns.

The application of rare events logistic regression analysis in modelling archaeological site patterns and probabilities faces several drawbacks, though, and requires caution with regard to their interpretation. As in any other type of regression analysis, careful definition of the dependent variable(s) and the appropriate spatial scale, ideally reflecting a spatially and thematically homogeneous set of localisation factors, is crucial for the success of the model but often difficult to assess a priori. Inaccuracies related to the dating and geographic localisation of archaeological sites can indeed bias the dependent variable and hence the model outcomes considerably, as well as the spatial non-uniformity of the sites' preservation and discovery potential and of archaeological prospection intensity. The former was partly addressed by the careful selection of only well-documented, high-quality sites (cf. section 2.2). However, comparison of observed burial and habitation site densities with the soil's erosion potential and modern agricultural land use in the studied regions confirms the spatially variable visibility and preservation potential of sites. The incompleteness of the archaeological site record and the associated uncertainty regarding the validity of negative observations, as well as the scarcity of positive observations in the dependent variable in general, further complicates both the establishment of a statistically solid regression relation and its evaluation, despite the application of rare events corrections. We are hence unable to quantify to which extent the modelled probability patterns reflect the true distribution of the Roman and Merovingian occupation. An alternative approach to deal with the lack of true site absence data, is the application of multiple cross-validation using Monte Carlo analyses. Finally, the selection of independent variables to be tested on their significance and influence in relation with the archaeological site pattern is equally decisive for the final model performance. Particularly the lack of spatially explicit and quantitative information on socioeconomic, political and cultural aspects of the Roman and Merovingian settlement system results in predominantly physical-deterministic probability models and constitutes a weakness of the current and analogous studies. The fact that our models do not yield perfect predictions of site locations is indeed an indication of the importance of the role that other factors constituted in ancient societies' settling choices. An integrated system approach that combines the presented models with dynamical theory-driven simulations might contribute to a better approach of such issues. Alternatively, empirical models can be improved by including additional explanatory variables that reflect measurable properties of the archaeological remains themselves, or spatial proxies of human perception of the environment (e.g. landscape visibility and accessibility; Wheatley 1995; Whitley 2002; Ridges 2006; Verhagen *et al.* 2013). More interdisciplinary studies need to be performed in order to examine the agricultural landscape, the husbandry and agricultural activities of the settlements and the local exchange networks in which they are involved.

Nevertheless, despite the risk of biases in the archaeological record and in the model setup, this study demonstrated that logistic regression models can significantly contribute to our understanding of the relation between past human occupation and the physical environment, and the functional, temporal and regional differentiation of these relations. Moreover, such models also provide a valuable means to spatially and statistically synthesise the pool of archaeological site data at the regional scale, as well as to test hypotheses regarding the relation between site locations and any spatial variable of interest.

Finally, the modelled site probability maps allow to assess likely patterns of past human activity that can be of interest in archaeological management and planning, but are also useful tools to include archaeological information in applications that require spatially continuous input data, including historical land cover modelling.

Acknowledgements

The work presented in this article is supported by the Interuniversity Attraction Poles programme of the Belgian Science Policy Office (research projects VI/22 and CORES VII/09), the KU Leuven Center for Archaeological Sciences, the Research Foundation - Flanders (FWO), and the Fund for Scientific Research - FNRS.

References

- Akaike, H. (1973): "Information theory and an extension of the maximum likelihood principle", in: *Second International Symposium on Information Theory*, Tsahkadsor, Armenia, 267-281.
- Allen, K. M. S., S. W. Green and E. B. W. Zubrow, ed. (1990): *Interpreting Space: GIS and Archaeology*, London.
- Andresen, J., T. Madsen and I. Scollar, ed. (1992): *Computing the Past: Computer Applications and Quantitative Methods in Archaeology*, Aarhus.
- Bai, S.-B., J. Wang, G.-N. Lü, P.-G. Zhou, S.-S. Hou and S.-N. Xu (2010): "GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China", *Geomorphology*, 115, 23-31.
- Binford, L. R. (1982): "The archaeology of place", *Journal of Anthropological Archaeology*, 1, 5-31.
- Bitter, P. (1991): "Romeinse tijd en vroege middeleeuwen: continuïteit en discontinuïteit", in: Bloemers van & Dorp, ed. 1991, 329-338.
- Bloemers van, J. H. F. and T. Dorp, ed. (1991): *Pre- en Protohistorie van de Lage Landen*, Houten.
- Bloemers, T., H. Kars, A. Van Der Valk and M. Wijnen, ed. (2010): *The Cultural Landscape & Heritage Paradox: Protection and Development of the Dutch Archaeological-Historical Landscape and Its European Dimension*, Amsterdam.
- Broothaerts, N., G. Verstraeten, C. Kasse, S. Bohncke, B. Notebaert and J. Vandenberghe (2014): "Reconstruction and semi-quantification of human impact in the Dijle catchment, central Belgium: a palynological and statistical approach", *Quaternary Science Reviews*, 102, 96-110.
- Brulard, T. (1960): *La Hesbaye. Étude géographique d'économie rurale*, PhD thesis, University of Leuven.
- Brulet, R., éd. (2008a): *Les Romains en Wallonie*, Brussels.
- Brulet, R. (2008b): "Fortifications de hauteur et habitat perché de l'Antiquité tardive au début du Haut Moyen-Âge, entre Fagne et Eifel", in: Steuer & Bierbrauer, ed. 2008, 13-70.
- Buck, C. E., W. G. Cavanagh and C. D. Litton (1996): *Bayesian Approach to Interpreting Archaeological Data*, Chichester.
- Bührer-Thierry, G. and C. Mériaux (2010): *La France avant la France. 481-888*, Paris.
- Butzer, K. W. (1982): *Archaeology as human ecology: method and theory for a contextual approach*, London.
- Canning, S. (2005): " 'Belief in the Past: Dempster-Shafer Theory, GIS and Archaeological Predictive Modelling", *Australian Archaeology*, 6-15.
- Cardot, F. (1987): *L'espace et le pouvoir. Étude sur l'Austrasie mérovingienne*, Histoire ancienne et médiévale 17, Paris.
- Carmichael, D. (1990): "GIS predictive modeling of prehistoric site distribution in central Montana", in: Allen *et al.*, ed. 1990, 216-225.
- Carr, C., ed. (1985): *For concordance in archaeological analysis: bridging data structure, quantitative technique, and theory*, Kansas.
- Childe, V. G. (1934): "Neolithic settlement in the west of Scotland", *Scottish Geographical Magazine*, 50, 18-25.
- Clarke, D. L., ed. (1972): *Models in archaeology*, London.
- Cohen, J. (1960): "A coefficient of agreement for nominal scales", *Educational and psychological measurement*, 20, 37-46.
- Conolly, J. and M. Lake (2006): *Geographical Information Systems in Archaeology*, London.
- Corbiau, M.-H. (2008): "La voie romaine Metz-Arlon-Tongres. Itinéraire, équipement et chronologie", in: Witvrouw & Gava, dir. 2008, 17-44.

- Corbiau, M.-H. and J.-M. Yante (2010): "Le réseau routier antique: adaptation, survie, déclassement au Moyen Âge. L'exemple d'itinéraires entre Arlon et la Meuse", in: Yante & Bultot-Verleysen, ed. 2010, 199-217.
- Crowley, L. (2011): "The role of mortuary ritual in the construction of social boundaries by privileged social groups within villa landscapes", in: Roymans & Derks, ed. 2011, 192-210.
- Crema, E. R., A. Bevan and M. W. Lake (2010): "A probabilistic framework for assessing spatio-temporal point patterns in the archaeological record", *Journal of Archaeological Science*, 37, 1118-1130.
- De Boe, G. (1971): "De stand van het onderzoek der Romeinse villa's in België", in: De Boe, ed. 1971, 5-14.
- De Boe, G., ed. (1971): *De stand van het onderzoek der Romeinse villa's in België / Une villa romaine à Haccourt (Liège). Rapport provisoire des fouilles 1967-1970*, Archaeologia Belgica 132, Brussels.
- De Clercq, W. (2009): *Lokale gemeenschappen in het Imperium Romanum: transformaties in de rurale bewoningsstructuur en de materiële cultuur in de landschappen van het noordelijk deel van de civitas Menapiorum (Provincia Gallia-Belgica, ca. 100 v. Chr.–400 n. Chr.)*, PhD thesis, Gent University.
- De Moor, G. and A. Pissart (1992): "Het reliëf", in: Denis, ed. 1992, 129-213.
- Defgnée, A. and A.-V. Munaut (1996a): "Evolution de l'environnement végétal du nord de la Gaule de la Tène à l'époque Gallo-Romaine", *RAPic*, 11, 325-331.
- (1996b): "Évolution du paysage au cours des occupations humaines sur le Grognon à Namur", in: Plumier, dir. 1996, 127-130.
- Denis, J., ed. (1992): *Geografie van België*, Brussels.
- Dikau, R. (1989): "The application of a digital relief model to landform analysis", in: Raper, ed. 1989, 51-77.
- Dupriez, G. L. and R. Sneyers (1978): *Les normales du réseau pluviométrique belge*, Brussels.
- Duvosquel, J.-M. and A. Dierkens, ed. (1991): *Villes et campagnes au Moyen Âge. Mélanges Georges Despy*, Centre belge d'histoire rurale. Publication, 97, Liège, 563-597.
- Ebert, J. I. (2000): "The state of the art in 'inductive' predictive modeling: seven big mistakes (and lots of smaller ones)", in: Westcott & Brandon, ed. 2000, 129-134.
- Ejstrud, B. (2005): "Taphonomic models: Using Dempster-Shafer theory to assess the quality of archaeological data and indicative models", in: Van Leusen & Kamermans, ed. 2005, 183-194.
- Ellison, A. and J. Harriss (1972): "Settlement and land use in the prehistory and early history of southern England: a study based on locational models", in: Clarke, ed. 1972, 91-962.
- Fawcett, T. (2006): "An introduction to ROC analysis", *Pattern recognition letters*, 27, 861-874.
- Ferdière, A. (2004): "Indigènes et "romanisés" à travers la tombe privilégiée en Gaule", *Latomus*, 63, 35-57.
- Ferdière, A., dir. (2004): *Capitales éphémères. Des capitales de cités perdent leur statut dans l'Antiquité tardive, Actes du colloque organisé par le Laboratoire Archéologie et Territoires, Tours, 6-8 mars 2003*, Revue Archéologique du Centre de la France, 25^e suppl., Tours.
- Finke, P., E. Meylemans and J. Van de Wauw (2008): "Mapping the possible occurrence of archaeological sites by Bayesian inference", *Journal of Archaeological Science*, 35, 2786-2796.
- Fox, C. (1943): *The personality of Britain: its influence on inhabitant and invader in prehistoric and early historic times*, Cardiff.
- Gauthier, N. (2002): "Tongres-Maastricht-Liège", in: Gauthier et al., ed. 2002, 71-73.
- Gauthier, N., B. Beaujard and F. Prévot, ed. (2002): *Topographie chrétienne des cités de la Gaule des origines au milieu du VIII^e siècle. XII. Province ecclésiastique de Cologne (Germania Secunda)*, Paris.
- Godart, M.-F. and C. Feltz (2009): *Atlas des Paysages de Wallonie. 2. Les Plateaux brabançon et hesbignon*, Namur.
- (2010): *Atlas des Paysages de Wallonie. 3. Le Plateau condrusien*, Namur.
- Guizard-Duchamp, F. (2003): "Fleuves, forêts et territoire dans les sources narratives des VI^e et VII^e siècles", *Revue du Nord*, 351, 573-594.
- Green, R. C. (1963): *A review of the prehistoric sequence of the Auckland Province*, Auckland.
- Grimes, W. F. (1945): "Early man and soils of Anglesey", *Antiquity*, 19, 169-174.
- Groot, M., D. Lentjes and J. Zeiler, ed. (2013): *Barely Surviving or More than Enough? The environmental archaeology of subsistence, specialisation and surplus food production*, Leiden.
- Hastie, T. and W. Fithian (2013): "Inference from presence-only data; the ongoing controversy", *Ecography*, 36, 864-867.
- Heymans, H. (1978): "De topografie van de Merovingische grafvelden in Belgisch Limburg en Maastricht", *Acta Archaeologica Lovaniensia*, 17, 66-134.
- Hollevoet, Y. (1991): "Een vroeg-middeleeuwse nederzetting aan de Hoge Dijken te Roksem (gem. Oudenburg)", *Archeologie in Vlaanderen*, 1, 181-196.
- (2016): "Entre Frisons, Francs et Anglo-Saxons : la Flandre maritime au haut Moyen Âge", in: Leroy & Verslype, ed. 2016, 69-76.
- Imai, K., G. King and O. Lau (2011): *Zelig: Everyone's Statistical Software. R package version 3.5.4*.

- Judge, W. J. and L. Sebastian, ed. (1988): *Quantifying the Present and Predicting the Past: Theory, Method and Application of Archaeological Predictive Modeling*, Denver.
- King, G. and L. Zeng (2001a): “Explaining rare events in international relations”, *International Organization*, 55, 693-715.
- (2001b): “Logistic regression in rare events data”, *Political analysis*, 9, 137-163.
- Kohler, T. A. and S. C. Parker (1986): “Predictive models for archaeological resource location”, *Advances in archaeological method and theory*, 9, 397-452.
- Kohler, T. A. (1988): “Predictive locational modelling: history and current practice”, in: Judge & Sebastian, ed. 1988, 19-59.
- Kooistra, L. I. (1996): *Borderland farming: possibilities and limitations of farming in the Roman period and early Middle Ages between the Rhine and Meuse*, Assen.
- Kooistra, L. I., L. van Beurden and K. Esser (2004): *De Romeinse villa van Kerkrade-Holzkuil onderzocht op organische resten*, Report BIAxiaal 176, Zaandam.
- Kvamme, K. (1985): “Determining empirical relationships between the natural environment and prehistoric site locations: a hunter-gatherer example”, in: Carr, ed. 1985, 208-238.
- Lefert, S. and K. Bausier (2013): “Villas gallo-romaines en Condroz namurois: des situations contrastées”, in: Vanmechelen, dir. 2013, 239-272.
- Lepot, A., A. Vallée, A. De Brue, B. Notebaert and G. Verstraeten (2016): *Le Condroz (Belgique) du I^{er} au VII^e siècle après J.-C.: Exploration du paysage rural antique*, in: Reddé, ed. 2016, 139-157.
- Leroy, I. and L. Verslype, ed. (2016): *Les cultures des littoraux au haut Moyen Âge. Cadres et modes de vie dans l'espace maritime Manche-mer du Nord du III^e au X^e s.*, Revue du Nord, Hors série. Collection Art et Archéologie, 24, Villeneuve-d'Ascq.
- Lock Stančić, G. Z., ed. (1995): *Archaeology and Geographical Information Systems: A European Perspective*, London.
- Maréchal, R. and J. Ameryckx (1992): “De bodems”, in: Denis, ed. 1992, 241-260.
- Massart, C. (1994): *Les tumulus gallo-romains conservés en Hesbaye: étude topographique*, Musées royaux d'art et d'histoire. Monographie d'archéologie nationale 9, Brussels.
- (2015): *Les Tumulus gallo-romains de Hesbaye (cité des Tongres). La représentation funéraire des élites*, Atuatuca 6, Tongres.
- Matterne, V. (2001): *Agriculture et alimentation végétale durant l'Âge du Fer et l'époque gallo-romaine en France septentrionale*, Archéologie des Plantes et des Animaux 1, Montagnac.
- Mehrer, M. W. and K. L. Wescott, ed. (2006): *GIS and Archaeological Site Location Modeling*, Boca Raton.
- Mertens, J. (1957): *Les routes romaines de la Belgique*, Brussels.
- Mithen, S.J., ed. (2000): *Hunter-Gatherer Landscape Archaeology: The Southern Hebrides Mesolithic Project, 1988-89*, Archaeological Fieldwork on Colonsay, Computer Modelling, Experimental Archaeology, and Final Interpretations, Cambridge.
- Morgan, J. M. and A. M. Lesh (2005): *Developing landform maps using Esri's ModelBuilder. 25th Esri International User Conference Proceedings*, San Diego.
- Munaut, A.-V. (1967): *Recherches paléo-écologiques en basse et moyenne Belgique*, Acta Geographica Lovaniensia 6, Leuven.
- Noël, R. (1991): “Moines et nature sauvage : dans l'Ardenne du haut Moyen Âge (saint Remacle à Cugnon et à Stavelot-Malmedy)”, in: Duvosquel & Dierkens, ed. 1991, 563-597.
- Noël, R., I. Paquay and J.-P. Sosson, ed. (2003): *Au-delà de l'écrit. Les hommes et leurs vécus matériels au Moyen Âge à la lumière des sciences et des techniques. Nouvelles perspectives, Actes du colloque international de Marche-en-Famenne, 16-20 octobre 2002*, Turnhout.
- Notebaert, B., G. Verstraeten, D. Vandenberghe, E. Marinova, J. Poesen and G. T. Govers (2011): “Changing hillslope and fluvial Holocene sediment dynamics in a Belgian loess catchment”, *Journal of Quaternary Science*, 26, 44-58.
- Nouwen, R. (2006): *De Romeinen in België (31 v.C. - 476 n.C.)*, Leuven.
- Opsteyn, L. and E. Taayke (1998): *De Frankische migratie. Nederland en Vlaanderen in de laat-Romeinse tijd*, Leuven-Amsterdam.
- Parker, S. C. (1985): “Predictive modeling of site settlement systems using multivariate logistics”, in: Carr, ed. 1985, 173-207.
- Peterson, C. E. and R. D. Drennan (2005): “Communities, settlements, sites, and surveys: regional-scale analysis of prehistoric human interaction”, *American Antiquity*, 70, 5-30.
- Phillips, S. J., M. Dudík, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick and S. Ferrier (2009): “Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data”, *Ecological Applications*, 19, 181-197.
- Pigière, F. (2009): *Évolution de l'économie alimentaire et des pratiques d'élevage de l'Antiquité au haut Moyen Âge en Gaule du nord. Une étude régionale sur la zone limoneuse de la Moyenne Belgique et du sud des Pays-Bas*, British Archaeological Reports-International Series 2035, Oxford.
- Pigière, F. and A. Lepot (2013): “Food production and exchanges in the civitas Tungrorum”. Autumn Meeting of the Association for Environmental Archaeology”, in: Groot et al., ed. 2013, 225-246.
- Pigière, F. (2015): “Animal husbandry in the core area of the civitas Tungrorum”, in: Roymans et al., ed. 2015, 163-175.
- Plumier, J. dir. (1996): *Cinq années d'archéologie en province de Namur 1990-1995*, Études et Documents. Fouilles, 3, Namur.
- R Core Team (2013): *R: A language and environment for statistical computing*, Vienna.
- Raepsaet, G. and M.-T. Raepsaet-Charlier (1975): “Gallia Belgica et Germania Inferior. Vingt-cinq années de recherches histo-

- riques et archéologiques”, in: Temporini & Haase, ed. 1975, 3-299.
- Raepsaet-Charlier, M.-T. (1994): “La cité des Tongres sous le Haut-Empire. Problèmes de géographie historique”, *Bonner Jahrbücher*, 194, 43-59.
- Raepsaet-Charlier, M.-T. and A. Vanderhoeven (2004): “Tongres au Bas-Empire romain”, in: Ferdière, dir. 2004, 51-73.
- Raepsaet-Charlier, M.-T. and G. Raepsaet (2011): “Villes et agglomérations de Belgique sous le Principat : les statuts”, *RBPh*, 89, 633-657.
- Raepsaet, G. (2013): “L’ethnogenèse de la *civitas Tungrorum* et la formation de la Province de Germanie”, *AC*, 82, 111-148.
- Raper, J. F., ed. (1989): *Three dimensional applications in geographical information systems*, London.
- Ridges, M. (2006): “Regional dynamics of hunting and gathering: an Australian case study using archaeological predictive modeling”, in: Mehrer & Wescott, ed. 2006, 123-143.
- Rogge, M. (1981): “Een Merovingische nederzetting te Avelgem - Kerkhove (West-Vlaanderen)”, in: Van Doorselaer, ed. 1981, 123-143.
- Reddé, M., ed. (2016): *Méthodes d’analyse des différents paysages ruraux dans le nord-est de la Gaule romaine* [en ligne], mis en ligne le 13 janvier 2016, consulté le 11 février 2017. URL : <https://hal.archives-ouvertes.fr/hal-01253470>
- Rommens, T., G. Verstraeten, L. Peeters, J. Poesen, G. Govers, A. Van Rompaey, B. Mauz, S. Packman and A. Lang (2007): “Reconstruction of late-Holocene slope and dry valley sediment dynamics in a Belgian loess environment”, *The Holocene*, 17, 777-788.
- Roosens, H. (1949): *De Merovingische begraafplaatsen in België. Repertorium en algemene beschouwingen*, Gent.
- Rose, M.R. and J. H. Altschul (1988): “An overview of statistical method and theory for quantitative model building”, in: Judge & Sebastian, ed. 1988, 173-255.
- Roymans, N. (1996): “The sword or the plough. Regional dynamics in the romanisation of Belgic Gaul and the Rhineland area”, in: Roymans, ed. 1996, 9-126.
- Roymans, N., ed. (1996): *From the Sword to the Plough. Three studies on the earliest romanisation of northern Gaul*, Amsterdam.
- Roymans, N. and T. Derks (2011): “Studying Roman villa landscapes in the 21st century. A multi-dimensional approach”, in: Roymans & Derks, ed. 2011, 1-44.
- Roymans, N. and T. Derks, ed. (2011): *Villa landscapes in the Roman North: economy, culture and lifestyles*, Amsterdam.
- Roymans, N., T. Derks and H. Hiddink, ed. (2015): *The Roman villa of Hoogeloon and the archaeology of the periphery*, Amsterdam.
- Sanders, W. T. (1956): “The Central Mexican symbiotic region: a study in prehistoric settlement patterns”, in: Willey, ed. 1956, 115-127.
- Steuer, H. and V. Bierbrauer, ed. (2008): *Höhensiedlungen zwischen Antike und Mittelalter von den Ardennen bis zur Adria*, Ergänzungsbände zum Reallexikon der Germanischen Altertumskunde, 58, Berlin-New York.
- Temporini, H. and W. Haase, ed. (1975): *Aufstieg und Niedergang der römischen Welt II*, Berlin.
- Van Den Eeckhaut, M., T. Vanwallegem, J. Poesen, G. Govers, G. Verstraeten and L. Vandekerckhove (2006): “Prediction of landslide susceptibility using rare events logistic regression: a case-study in the Flemish Ardennes (Belgium)”, *Geomorphology*, 76, 392-410.
- Van Doorselaer, A., ed. (1981): *De Merovingische beschaving in de Scheldevallei*, Kortrijk.
- Van Oost, K., G. Govers and P. Desmet (2000): “Evaluating the effects of changes in landscape structure on soil erosion by water and tillage”, *Landscape Ecology*, 15, 577-589.
- Van Ossel, P. (1983): “L’établissement romain de Loën à Lixhe et l’occupation rurale au Bas-Empire dans la Hesbaye Liégeoise”, *Helinium*, XXIII, 143-169.
- (1992): *Établissements ruraux de l’Antiquité tardive dans le nord de la Gaule*, Gallia Suppl. 51, Paris.
- Van Ossel, P. and P. Ouzoulias (2000): “Rural settlement economy in Northern Gaul in the Late Empire: an overview and assessment”, *JRA*, 13, 133-160.
- Van Ossel, P. (2010): “De la ‘villa’ au village : les prémices d’une mutation”, in: Yante & Bultot-Verleysen, ed. 2010, 219-236.
- Van Leusen, M. and H. Kamermans, ed. (2005): *Predictive Modelling for Archaeological Heritage Management: A Research Agenda*, Amersfoort.
- Van Rompaey, A., G. Verstraeten, K. Van Oost, G. Govers and J. Poesen (2001): “Modelling mean annual sediment yield using a distributed approach”, *Earth Surface Processes and Landforms*, 26, 1221-1236.
- Vanacker, V., G. Govers, P. van Peer, C. Verbeek, J. Desmet and J. Reyniers (2001): “Using Monte Carlo simulation for the environmental analysis of small archaeological datasets, with the Mesolithic in northeast Belgium as a case study”, *Journal of Archaeological Science*, 28, 661-669.
- Vanderhoeven, A. (1996): “The earliest urbanisation in Northern Gaul : some implications of recent research in Tongres”, in: Roymans, ed. 1996, 189-260.
- (2002): “Tongres”, in: Gauthier *et al.*, ed. 2002, 75-89.
- Vanderhoeven, A. and A. Erynck (2007): “Not in my back yard? The industry of secondary animal products within the Roman civitas capital of Tongeren (Belgium)”, in: *Promoting Roman finds: context and theory*, Oxford, 156-175.
- Vanmechelen, R., dir. (2013): *Archéologie entre Meuse et Hoyoux. 2. Contexte, analyses*, De la Meuse à l’Ardenne, 45, Saint-Hubert.

- Vanpoucke, S., F. Pigière, A. Defgnée and W. Van Neer (2007): "Pig husbandry and environmental conditions in northern Gaul during Antiquity and the early Middle Ages: the contribution of hypoplasia analysis", *Archaeofauna*, 16, 7-20.
- Vanwalleghe, T., H. Bork, J. Poesen, M. Dotterweich, G. Schmidtchen, J. Deckers, S. Scheers and M. Martens (2006): "Prehistoric and Roman gullying in the European loess belt: a case study from central Belgium", *The Holocene*, 16, 393-401.
- Vanwalleghe, T., M. Van Den Eeckhaut, J. Poesen, G. Govers and J. Deckers (2008), "Spatial analysis of factors controlling the presence of closed depressions and gullies under forest: Application of rare event logistic regression", *Geomorphology*, 95, 504-517.
- Verhagen, P. (2006): "Quantifying the qualified: The use of multicriteria methods and Bayesian statistics for the development of archaeological predictive models", in: Mehrer & Wescott, ed. 2006, 191-216.
- Verhagen, P., H. Kamermans, M. van Leusen and B. Ducke (2010): "New developments in archaeological predictive modelling", in: Bloemers *et al.*, ed. 2010, 431-444.
- Verhagen, P., L. Nuninger, F.-P. Tourneux and K. Jeneson (2013): "Introducing the human factor in predictive modelling: A work in progress", in: *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*, Amsterdam.
- Verslype, L. (2003a): "À la vie, à la mort. Considérations sur l'archéologie et l'histoire des espaces politiques, sociaux et familiaux mérovingiens", in: Noël *et al.*, ed. 2003, 405-460.
- (2003b): " 'Limites sans frontières'. Réflexions sur la perception archéologique de l'espace en Neustrie septentrionale (V^e-VIII^e siècles)", *Revue du Nord*, 3-351, 551-572.
- (2010): "Avant le village, l'habitat rural mérovingien. Les sources archéologiques au nord-ouest des royaumes francs (fin V^e-VIII^e siècles)", in: Yante & Bultot-Verleysen, ed. 2010, 237-275.
- (2013): "Les établissements ruraux du Haut Moyen Âge : nécropoles et indices d'occupation en Condroz", in: Vanmechelen, dir. 2013, 273-300.
- Verstraeten, G. (2006): "Regional scale modelling of hillslope sediment delivery with SRTM elevation data", *Geomorphology*, 81, 128-140.
- Verstraeten, G., K. Van Oost, A. Van Rompaey, J. Poesen and G. Govers (2002): "Evaluating an integrated approach to catchment management to reduce soil loss and sediment pollution through modelling", *Soil Use and Management*, 18, 386-394.
- Verstraeten, G., A. Van Rompaey, J. Poesen, K. Van Oost, G. Govers and L. Stalpaert (2003): "Thema 2.20 Kwaliteit Bodem: erosie. MIRA-T 2003", *Milieu en Natuur rapport Vlaanderen*, 345-355.
- Warren, R. E. and D. L. Asch (2000): "A predictive model of archaeological site location in the eastern Prairie Peninsula", in: Westcott & Brandon, ed. 2000, 6-36.
- Westcott, K. J. and R. J. Brandon, ed. (2000): *Practical Applications of GIS For Archaeologists. A Predictive Modeling Kit*, London.
- Wheatley, D. (1992): "Going over old ground: GIS, archaeological theory and the act of perception", in: Andresen *et al.*, ed. 1992, 133-138.
- (1995): "Cumulative viewshed analysis: a GIS-based method for investigating intervisibility, and its archaeological applications", in: Lock Stančič, ed. 1995, 171-185.
- Whitley, T. G. (2002): "Spatial variables as proxies for modeling cognition and decision-making in archaeological settings: A theoretical perspective", in: *24th Annual Meeting of the Theoretical Archaeology Group. Manchester*.
- Willey, G. R., ed. (1956): *Prehistoric settlement patterns in the New World*, New York.
- Witvrouw, J. and G. Gava, dir. (2008): *Le pont romain et le franchissement de la Meuse à Amay. Archéologie et Histoire*, Bulletin du Cercle Archéologique Hesbaye-Condruz, 29, Amay.
- Woodman, P. E. (2000): "A predictive model for Mesolithic site location on Islay using logistic regression and GIS", in: Mithen, ed. 2000, 445-464.
- Yante, J.-M. and A. M. Bultot-Verleysen, ed. (2010): *Autour du "village". Établissements humains, finages et communautés rurales entre Seine et Rhin (IV^e-XIII^e siècles)*, Actes du colloque international de Louvain-la-Neuve, 16-17 mai 2003, Publications de l'Institut d'études médiévales 25, Louvain-la-Neuve.