Archeologia e Calcolatori
25, 2014, 223-239

A REVIEW OF CASE STUDIES IN ARCHAEOLOGICAL LEAST-COST ANALYSIS

1. Introduction

In recent years, the number of archaeological studies applying least-cost analysis (LCA) increased fairly constantly, culminating in the book by White and Surface-Evans (2012). This contribution is to discuss not only the case studies of the book just mentioned but also several other recent publications on this subject from a technical point of view. Due to space limitations only a selection of recent case studies could be analysed. The aim of the case studies considered is to calculate site catchments, least-cost paths, least-cost networks, or accessibility. In the first part, the methodology of the case studies with respect to these aims is discussed.

The second part deals with assessing the reliability of the results obtained in the case studies. This assessment is based on several aspects relevant for all LCA applications:

– accuracy and resolution of the geographical data used including a discussion of landscape change since the period considered;
– the cost model assigning friction to movements within the landscape;
– LCA software;
– varying parameters in the cost model to analyse the stability of the results;
– validation of the results.

An overview of the case studies and their assessment is presented in Table 1. Some of the concepts discussed are illustrated by examples from a small hilly study area in Germany.

2. Case studies: site catchments

Probably the most basic application of LCA is the generation of a least-cost site catchment (LCSC). The technique for creating a LCSC by GIS is about as old as least-cost path calculations (for instance Gaffney, Stančič 1992). According to Conolly and Lake (2006, 214), site catchment analysis «is an investigation of the resources available within a region (catchment) accessible from a site». In general, a LCSC includes all areas that can be reached by expending less than a preset cost limit. When costs are measured in terms of time, the boundary of the site catchment area is often called an isochrone. Fig. 1a shows examples of LCSCs for three different cost models.

Case study	DEM	Software	Cost function s = slope	Features
Hudson 2012	10 m	ArcGIS Spatial Analyst: Distance Cost Weighted tool, Distance Shortest Path tool	ArcGIS default settings	
Kantner 2012	10 m 30 m	ArcGIS's Costdistance	Slope, Tobler, Pandolf *et al.*	V
Kienlin, Cappenberg, Korczynska 2012	21 m	ArcGIS 9.3 Path Distance	Tobler(s) = $6e^{-3.5\|s+0.05\|}$ with s = vertical change divided by horizontal change	
Livingood 2012	180 m	Custom software: Dijkstra and A*	Tobler combined with canoe-travel	A
Murrieta-Flores 2012	10 m	IDRISI TAIGA	cost(s) = $0.031s^2 - 0.025s + 1$ with s = slope, in degree; additional layer with intuitive costs for crossing rivers	G
Nolan, Cook 2012	20 m	ArcGIS 9.3 spatial analyst, Distance, Shortest Path	ArcGIS 9.3 default settings	(G)
Phillips, Leckman 2012	10 m	ArcGIS Path Distance	Tobler (s)	AG(P)
Posluschny 2013		ArcGIS and custom software	ArcGIS default settings and critical slope (12%) function	
Rademaker, Reid, Bromley 2012	90 m	ArcGIS Path Distance	Tobler (s), but mainly Pandolf (s): $1.5W + 2.0(W+L)(L/W)^2$ $+ N(W+L)(1.5V^2 + 0.35Vs)$ with s = slope, in percent W = weight of walker (50 or 70 kg) L = load (0 or 25 kg) V = walking speed (0.4, 1.2 or 2.4 m/s) N = terrain coefficient (default: 1)	G(P)V
Richards-Rissetto 2012	?	?	Values assigned to facilitators and barriers combined with slope	
Rissetto 2012	90 m	ArcGIS 9.2 spatial analyst	ArcGIS default settings	(P)
Surface-Evans 2012	10 m	ArcGIS spatial analyst: Cost Weighted Distance, Corridor function	ArcGIS default settings and Tobler (s)	V
Ullah, Bergin 2012	5 m	GRASS r.walk	Langmuir : $a{*}\Delta_1 + b{*}\Delta_2 + c{*}\Delta_3 + d{*}\Delta_4$ with Δ_1 = distance covered Δ_2 = uphill altitude difference Δ_3 = moderately downhill slope: altitude difference Δ_4 = steep downhill slope: altitude difference	P
Verhagen, Jeneson 2012	5 m 35 m	ArcGIS Path Distance	Tobler (s) combined with openness	(P)V
White 2012	30 m	Custom software combined with ESRI's Network Analyst	Refined Pandolf (s) for s > 0: V estimated using Tobler(s) W = 60 kg; L = 5 kg N = 1.2 in vegetated areas or loose soil additional formula for s < 0	G

Tab. 1 – Overview over the case studies discussed; the DEM entry shows the cell size of the DEM; the following abbreviations for features are used: A = Evidence of anisotropic calculation (return paths), G = Ground truthing of the LCA results; P = Discussion of palaeogeography, i.e. changes in the physical landscape since the period considered; (P) if palaeogeography is discussed but calculations apply unmodified modern DEM; V = Varying parameters in the cost model to assess the stability of the LCA results.

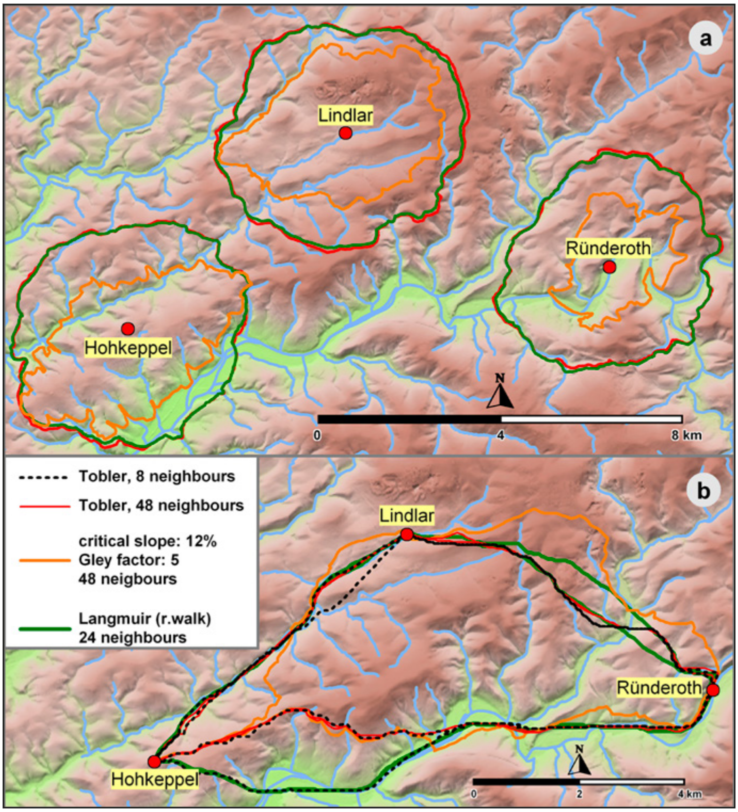

Fig. 1 – (a) Site catchments and (b) LCPs applying anisotropic LCA algorithms and some of the cost functions depicted in Fig. 3.

A standard raster GIS procedure is to create an accumulated cost-surface (ACS) by spreading from the origin and storing the accumulated costs of movement for each raster cell. If the spreading process is stopped at a preset cost limit, a LCSC results. The outcome of a site catchment analysis for a given site location is dependent on three choices: (i) the predefined cost limit, (ii) the cost model, and (iii) the spreading algorithm. Cost models and spreading algorithms are discussed in more detail below.

Kienlin, Cappenberg and Korczynska (2012) calculate site catchments delimited by 15 minute isochrones for two late Bronze Age sites in southern Poland with a straight-line distance of about 700 m. They apply Tobler's slope-dependent cost function for estimating walking time and ESRI's Path Distance tool provides the spreading procedure. The authors briefly discuss the fact that the two 15 minute isochrone site catchments overlap and come to the conclusion that smaller site catchments might be sufficient to ensure adequate harvests for the settlements considered. If the two sites are contemporary, a method combining LCSC with least-cost Voronois (Herzog 2013c) seems to be more appropriate.

Ullah and Bergin (2012) calculate LCSCs in their case study applying agent-based simulations for modelling the impact of the economy of villages in eastern Spain on their environment. For each village, a LCSC is computed using the GRASS GIS procedure r.walk with the default cost function estimating walking time. Four hypothetical village location strategies are investigated within this framework: optimal access to farm land and pastures, optimum seclusion, optimum defensibility, and optimum viewshed. This approach allows to watch the effects of different parameters and is a move towards assessing the impact of dynamic processes. But with respect to LCA the model is kept static because catchment calculations are computationally intensive. In the model, the land value of slopes decreases only if the gradient exceeds 10° (i.e. 17.5%). But according to the cost function used by Ullah and Bergin, walking 100 m on flat terrain requires 72 seconds, whereas ascending a slope of 15% covering the same distance takes 162 seconds. Unfortunately, the cost surfaces generated do not include the costs of returning to the village and consequently, «households avoided using high-slope areas only if they had to walk up the slope».

A non-traditional LCSC analysis is presented in the case study by Sarah Surface-Evans (2012). Her aim is to test the model that Shell Mound Archaic (SMA) «sites represent a shift from a foraging strategy to a collector strategy». She applies the ESRI cost corridor function for all SMA sites in the study area resulting in a raster that is twice the sum of the ACSs of these locations. The cells with the lowest 10% of the values in this raster form the areas designated as site catchments. In fact, this approach generates a least-cost point density map, the "catchments" are the areas with highest point density. So a group of nearby sites will be assigned a large common catchment, whereas only a very small catchment is generated for a site located far from all others. Extending the study area will probably change the catchment sizes. Omitting a site from the calculation will have an impact on the catchments of all other sites. Of course, such an impact is desired if the aim is to avoid overlapping catchment zones, but this aim is not achieved. Due to these observations, this alternative approach to LCSC is not recommended.

3. Case studies: least-cost paths

Probably the most popular application of LCA in archaeology is the generation of a least-cost paths (LCPs) connecting a set of site locations. This task is closely related to the calculation of least-cost networks, but reconstructing networks is more complex and discussed in a separate chapter below. The LCP between two locations is the route involving minimal costs in terms of the cost model selected. With the Dijkstra algorithm, the LCP is derived from the ACS by backtracking. In the spreading process, only the most efficient steps are taken, and each step from a raster cell to one of its neighbours is recorded so that it is possible to retrace the steps to the origin.

An isotropic cost model assumes that costs are independent of the direction of travel, whereas anisotropic costs take the direction of travel within a raster cell into account (Conolly, Lake 2006, 215). Typically slope is anisotropic. The effort of traversing a cell with a given slope value depends on the direction: climbing the slope often involves higher costs than descending; if you walk on a path along the elevation contour line, the effort corresponds to that for walking on level ground, independent of the steepness of the slope to your left and right. With anisotropic LCP calculations, the resulting paths connecting two locations often differ depending on the direction (Fig. 1b). Alternatively, you may choose to use the same path for returning and therefore average the costs of movements in both directions. Such cost models were applied for calculating the LCSCs in Fig. 1a.

Archaeologists often calculate LCPs in order to reconstruct old routes. For instance, Verhagen and Jeneson (2012) try to reconstruct a 7 km stretch of the Roman road *Via Belgica* in the hilly Dutch region known as Limburg. The aim of Rademaker, Reid and Bromley (2012) is to reconstruct paths between coastal Palaeoindian sites and obsidian sources in Southern Highland Peru. Rissetto (2012) calculates LCPs to reconstruct the chert procurement of hunter-gatherers during the Magdalenian period (17,000-11,000 BP) in Northern Spain. Posluschny (2012) compares LCPs to route reconstructions created by traditional methods.

The case study by Livingood (2012) investigates the spatial pattern of flat-top pyramidal mounds (1000-1600 AD) in the southern Appalachians assuming that these mounds are located at civic-ceremonial centres of chiefdoms. The histogram of least-cost distances between the mounds shows two modes: the mode at the larger distance is considered to correspond to distances between mounds in different polities, and the smaller mode possibly reflects the typical distance of a secondary centre to its administrative centre. Livingood compares the distribution of Euclidian distances with that of the anisotropic least-cost distances. This comparison is only valid if the samples (i.e. distances) are independent. Therefore, calculating the average of the two least-cost distances between two locations would have been more appropriate so that the sample size for both sample sets is identical. This data could be used for an attempt to reconstruct the polity boundaries after identifying the primary centres (with an approach similar to that of Hare 2004).

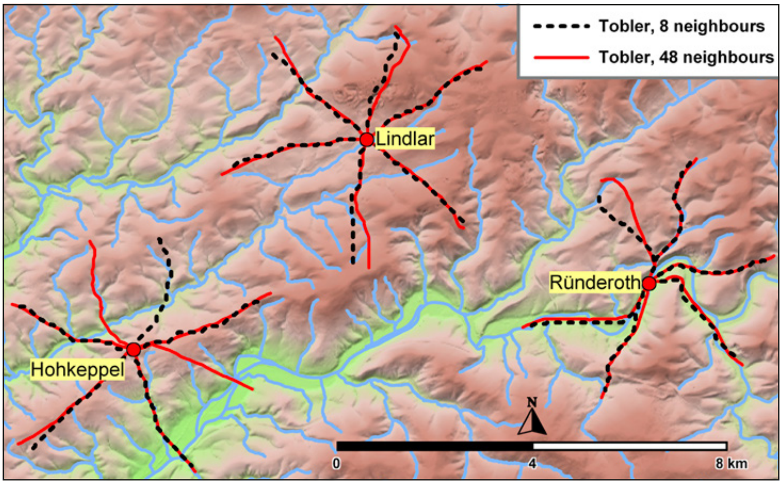

Fig. 2 – Focal networks.

Phillips and Leckman (2012) attempt to model water transport in the desert of New Mexico. A 10-m grid DEM forms the basis of the anisotropic calculations applying the Tobler hiking function. This contribution shows widely different paths for both directions (A to B and B to A) due to anisotropic computations. Selecting appropriate end points of the paths to be reconstructed is an issue in this case study because in one part of the study area only the origin (source) is known but no destinations. Generating focal networks (Fig. 2) based on a given origin as proposed by Fábrega Álvarez and Parcero Oubiña (2007) might have been more appropriate.

4. Case studies: least-cost networks

Posluschny (2012) wants to refute the hypothesis that the princely site called Glauberg played a major role in the Early Iron Age route network. Therefore he applies ArcGIS standard procedures to connect the sites with a certain type of Early Iron Age pottery by a least-cost network. The resulting route network avoids the Glauberg. The aim of Kantner's case study (2012) is to reconstruct paths between contemporary sites in Northwestern New Mexico and he, too, applies ArcGIS procedures to calculate the path network.

Different models exist for network generation (Herzog 2013c), and it is not quite clear which model is implemented in the GIS software used. For example, if the effort of constructing a route is high and people leave the vicinity of their house only rarely, a network with a small total of all route lengths results. This seems to be the objective of the networks shown in the case studies of Posluschny and Kantner. In contrast, if no construction effort is needed to create a path to a neighbour, most probably the direct path is taken resulting in a network with a large total path length. Moreover, a main road passing through the area before the first houses were built has some impact on the layout of the route network. Many other factors may play a role as well, like the sequence of building the houses or the size of the sites.

White (2012) applies the n nearest neighbour method to reconstruct road networks in the deserts of Arizona. Based on his tests with several values of n, White decides that connecting each location with the five nearest neighbours produces the best outcomes. The network reconstructions consist of several unconnected components, and often the shortest path in the network between two nearby locations is a lot longer than the direct shortest path. If the objective is to reconstruct the paths to the nearest neighbours, an approach based on least-cost triangulation avoids the drawbacks of the n nearest neighbour method (Herzog 2013c).

Nolan and Cook (2012) aim to reconstruct the trade relationships in Ohio during different time slices in the period from 801 to 1450 AD. In a first step, the authors create focal networks for each site by connecting the focal site to each other site within a 40 km circular catchment. These networks are compared to a set of focal networks where the targets are selected so that the paths from the focal point to the targets are most likely to be travelled, i.e. the progress on these paths is most cost-efficient. This is performed by subdividing each catchment (circular catchments with the radius in the range of 5 to 40 km are considered) into eight segments corresponding to eight directions (N, N-E, E, etc. like eight cake pieces), and calculating the most efficient path in each segment. Changing the orientation of the coordinate system by a few degrees will change the focal network as well: the examples in the case study show that this approach often generates unrealistic paths ending at the segment boundaries. This drawback is avoided by the focal network method (Fig. 2) proposed by Fábrega Álvarez and Parcero Oubiña (2007); moreover, it is based on LCSCs rather than circular catchments. Nolan and Cook also generate cost surfaces from the estimated harvest return differences. A more straight-forward approach might be more appropriate: estimate the amount of calories that must be imported by trade and find the most efficient paths to calories surplus locations in the vicinity so that the energy gap is filled.

5. LCA methods for calculating accessibility

The case study of Hudson (2012) deals with settlement patterns of two time periods (950-1150 AD and 1300-1600 AD) in a study area located in New Mexico. Accessibility of a settlement is measured by counting the number of different paths to this site, a method often applied in space syntax.

Four source points outside the study area were chosen to simulate people from outside moving into the area and the LCP to each site was calculated. In terms of space syntax, the accessibility of large sites is low, therefore Hudson concludes that «it may not be possible to interpret the result of landscape space syntax along the same line as traditional space syntax».

Richards-Rissetto (2012) uses another space syntax concept for calculating accessibility in her case study investigating the relationship between site configuration and social connectivity at the Maya site of Copán, Honduras in the period 763-820 AD: least-cost integration is the average cost of travel from one source point to all target points within a certain group. Such an approach is only valid if all sites in the study area are known. Richards-Rissetto's data is based on a full-coverage survey of the study area, but new sites detected just outside the study area might change the results of her tests.

The aim of Murrieta-Flores (2012) is to analyse the site location of prehistoric herding societies with respect to natural terrain accessibility in the mountainous area known as Sierra Morena in Spain. First, the access points at the border of the study area are identified by calculating pass locations on several scales using a morphometric approach. Each pair of access points is connected by a LCP with a cost model taking slope and the impediment of crossing certain types of rivers into account. Zones of high accessibility are calculated by counting for each cell the number of LCPs in its circular catchment area with a radius of 1.5 km^2. The application of a Kolmogorov-Smirnov test shows that the known prehistoric sites are closer to these high accessibility zones than random locations. Straight-line distances form the basis of the Kolmogorov-Smirnov test. According to Fig. 9 in that publication, Early Bronze Age monuments are clustered which could invalidate the test.

Alternative methods for calculating accessibility in a landscape based on traversal costs are probably more appropriate for most archaeological purposes (for an overview see Herzog 2013d). Several of these methods are not based on a route network but directly reflect the impact of the cost model on the landscape.

6. DEM and palaeogeography

Two aspects determine the quality of a digital elevation model (DEM): cell size and accuracy (Herzog 2014a). According to Table 1, the case studies used different cell sizes. Often the DEM cell size depends on data availability rather than on the requirements of the model. It is quite obvious that the elevation data with a 90 m cell size used by Rademaker, Reid, and Bromley (2012) is not able to represent a "steep-walled canyon" with a "~10-m-wide slot" properly.

Moreover, computation times increase considerably when using high DEM resolutions, especially with large study areas and if a "brute force" implementation of the spreading algorithm is used. Therefore, LIVINGOOD (2012) coarsened the DEM resolution of 30 m to "180-m² blocks" (which probably means 180 m by 180 m, i.e. 32,400 m²). Generally, the proportion of steep-slope cells decreases with increasing the cell size (SMITH, GOODCHILD, LONGLEY 2007, 260-261), though KANTNER (2012) maintains that many more low- or no-slope cells are present in a DEM when increasing the resolution.

Accuracy of the DEM is discussed in the case study of RADEMAKER, REID, and BROMLEY (2012). They had detected that the ASTER data is less reliable than the SRTM data and therefore chose the DEM with lower resolution. VERHAGEN and JENESON (2012) first used a DEM (cell size: 5 m) derived from LiDAR data. They found that the resulting LCP coincided partly with the modern motorway. For this reason, they used the ASTER DEM (cell size: approximately 35 m by 35 m) instead, with elevation in metres, not in centimetres, claiming that this is a better approximation of the relief in Roman times.

In the study area of MURRIETA-FLORES (2012), «quick and deep processes of erosion» occur. But she does not include this effect in her model.

To account for changes since the Bronze Age, KIENLIN, CAPPENBERG and KORCZYNSKA (2012) digitised water bodies from a historic map created at the end of the 18th century, but changes in landscape relief are not addressed.

RISSETTO (2012) mentions that the «position of the steep northern coastline has changed little – receding only 4 to 12 km to the south». A substantial proportion of his reconstructed paths runs along the modern northern coastline.

ULLAH and BERGIN (2012) apply a high resolution DEM grid (cell size: 5 m) and include a refined model of erosion-deposition phenomena in their agent-based simulations. But it seems that the start configuration of the model is the modern landscape, i.e. the accumulated result of these processes.

7. COST MODELS

Costs of movement in the landscape form the basis of each LCA study. A cost function is required that allows calculating the cost of each move from a raster cell to its neighbour. HUDSON (2012) thinks that the actual cost of travelling is of secondary importance for her analysis. Along similar lines, VERHAGEN and JENESON (2012) state that the differences between the Tobler and other hiking functions are not that big from a practical point of view. But case studies like that of KANTNER (2012) show that different cost models often produce widely different LCA results.

Many of the LCA studies discussed in this contribution consider slope-based costs only (Table 1). Two slope-based cost functions are very popular and used by many studies: the default cost function in ESRI software, which is simply slope (SURFACE-EVANS 2012), and the Tobler hiking function. Slope can be measured in degrees and percent, resulting in different cost functions, but some authors do not mention which unit of measurement was used (for instance Rissetto 2012; Nolan, Cook 2012). Many authors start with assigning costs proportional to slope because this is the easiest way to create a result with most LCA software. However, none of the case studies using this easy push-button option presents evidence that this cost function is appropriate. Kantner (2012) explains that slope is «a terrible proxy of movement costs across flat or nearly flat terrain» because flat surfaces have a slope of zero (Herzog 2013a, 2014a).

The Tobler cost function (Fig. 3) is more appropriate though not based on sound scientific data (Herzog 2013a). But the main disadvantage of the Tobler cost function is the danger of misunderstanding the formula. The formula requires slope measured in vertical change divided by horizontal change, which is neither percent nor degree slope, but percent slope divided by 100 – though Conolly and Lake (2006, 219) claim that slope in degrees is needed for this formula.

In general, the Tobler hiking function generates fairly direct LCPs (for instance Verhagen, Jeneson 2012) except on very steep slopes. But the case study in New Mexico (Phillips, Leckman 2012) presents paths with long detours and small-scale curves in the Doña Ana part of the study area. According to Google Earth, there are no large differences in elevation in this area. If the Tobler hiking function was not applied correctly, i.e. slope was overestimated by a factor of 100, this could have resulted in the long and wriggly paths.

Surface-Evans (2012) applies the Tobler formula and notes that average travel cost is 4.8 km/hr. This indicates that the Tobler function was not applied correctly: the function returns an estimate of the speed, which is high on flat terrain and low on steep slopes. To serve as a cost function, the Tobler formula has to be converted to a time estimate, resulting in low time requirements for covering a distance on level ground and high values for steep slopes. Surface-Evans herself describes the effect of applying the Tobler function wrongly: «In some ways, Tobler's hiking function excessively minimizes the impact of landscape travel barriers».

With any of these two misunderstandings in applying the Tobler formula, the LCA results and conclusions derived on this basis are seriously incorrect.

Ullah and Bergin (2012) apply the Langmuir cost function (Fig. 3) estimating time to calculate the cost surfaces. This rule of thumb function is not intuitive because it includes a large jump in costs at a downhill slope of 21.25% (Herzog 2013a).

Murrietta-Flores (2012) uses a binomial function (Fig. 3) derived from two values: the average walking speed on flat surfaces (5 km/hr) and in the Middle Mountain area of Nepal (384 m/hr).

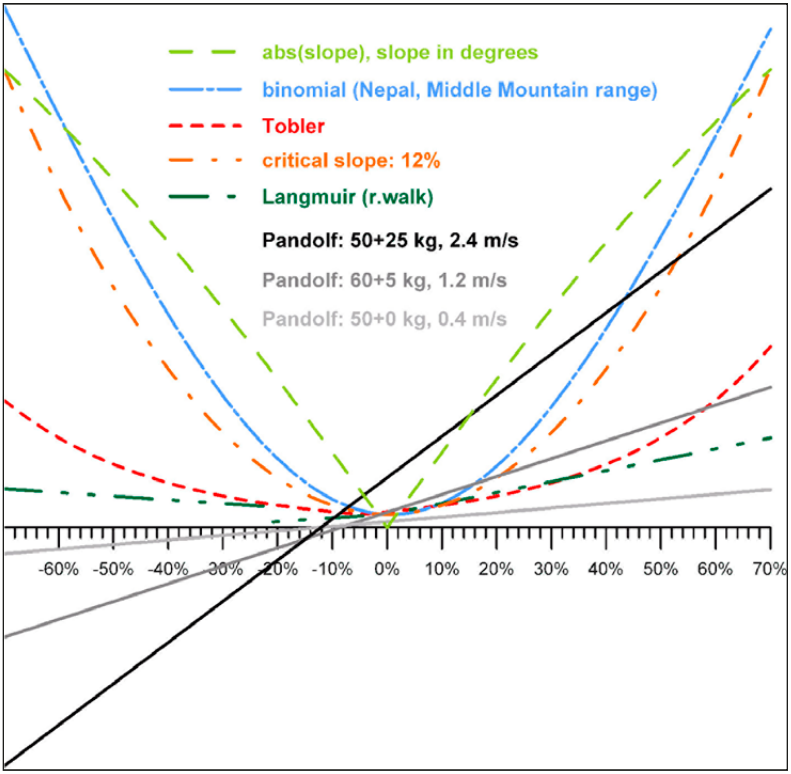

Fig. 3 – Comparison of slope-dependent cost functions used in the case studies. The cost functions are scaled to assist comparisons. The Pandolf cost functions are negative for steep downhill slopes.

The aim of the cost functions discussed above is to estimate travel time. An alternative approach is to minimize energetic expenditures. Often one of the currencies is chosen without discussing the advantages and disadvantages of the two approaches.

Livingood (2012) prefers cost models estimating time rather than energy expenditure because obviously historic accounts of journeys refer to time rather than the calories required. Kantner (2012) also chooses optimising travel time because his study area is a high desert without many sources of water. But there are some indications that a shortage of food was quite common in prehistoric periods, whereas people worked less than in modern times (e.g. Kerig 2008).

In their LCP calculations, Rademaker, Reid, and Bromley (2012) apply Pandolf's formula derived from measurements of energy consumption for uphill slopes in the range of 0 to 12%. This formula is slope-dependent but also includes parameters for the speed, weight and load of the walker and a terrain factor (Fig. 3). Rademaker and his colleagues assume that pedestrians walk at a constant speed irrespective of slope and terrain. White (2012) notes with reference to Pandolf's formula: «Walking downhill produces significantly negative cost estimates, which suggests that a traveler could recharge by walking in this direction […]. Since downhill walking still requires effort, the model is not entirely useful in this situation».

White (2012) applies Pandolf's cost function for ascending terrain, and another formula for downhill travel avoiding negative cost values. The speed parameter of this combined cost function is estimated using Tobler's formula, and a terrain factor of 1.2 is chosen for areas of loose soil or vegetation.

Verhagen and Jeneson (2012) apply the Tobler formula for reconstructing a Roman road though they observe that «Roman roads almost never take slopes over 15%». So a cost function for wheeled vehicles with a critical slope of 15% or less is more appropriate than a hiking function. Posluschny (2012) applies such a slope-dependent cost function with a critical slope of 12% (Herzog 2013a, 2014a).

In some case studies, authors are aware that other cost factors beyond slope might play a role as well. For many landscapes, river-based travel provides an alternative to land-based travelling, and creeks and rivers often are obstacles for walkers. This is an issue in several case studies using a slope-dependent cost model (Kienlin, Cappenberg, Korczynska 2012; Nolan, Cook 2012; Surface-Evans 2012). For instance, the Great Miami River with a breadth of approximately 90 m is running through the study area of Nolan and Cook (2012). The authors admit that river travel may have been important as well but that this was beyond the scope of their investigation. Discussing their results, they come to the conclusion that «not surprisingly, the lowest cost is incurred if one moves along the paths of extant waterways». But people could not walk on water but had to change to some sort of boat, and crossing the rivers might not have been easy. Kienlin, Cappenberg and Korczynska (2012) simply cut off the site catchments at the border of the flood plain of the Dunajec river.

The study area of Posluschny (2012) for the most part consists of hilly terrain with mostly small streams. He discusses the hypothesis that Iron Age routes preferred ridges to avoid wet flood plains and crossing tributaries in such an area. He presents archaeological evidence for built path structures in wet valleys and observes that the routes for long-distance transport of special goods often used rivers. In his view, the long-distance routes were only used during the appropriate season. So Posluschny thinks that slope is the main factor determining the Iron Age route layout in his study area.

The cost model of Livingood (2012) allows both overland and canoe travel. Overland travel combines Tobler's cost function with time penalties for crossing waterways or switching to canoe travel; the time needed for canoe travel is estimated on the basis of detailed data both from historic and modern accounts and adjusted according to the current.

As Murrieta-Flores (2012) notes, «topography is far from being the only influential factor in human movement», but modelling topography adequately is a good starting point. It is a pity that many of the case studies published so far did not reach this starting point.

8. LCA software

LCA software is available for isotropic and anisotropic calculations. Nearly all archaeological case studies performing LCA calculate costs on the basis of slope, so anisotropic approaches are more appropriate. Implementing software for anisotropic calculations is far more complex, and in general, the user interface is more complex, too. For example, the anisotropic ESRI ArcGIS procedure Path Distance supports four different input raster layers and several additional input parameters.

In general, LCA algorithms like Dijkstra or A* are defined for graphs, and the raster to vector conversion often creates problems if the number of links to neighbouring raster cells is low or if long links are generated that allow to skip over high cost barriers unnoticed (Herzog 2013b, 2014a). This raster to graph conversion is part of standard GIS LCA procedures. White (2012) creates this graph in his custom software (considering eight neighbouring cells) and uses the graph procedure provided in ESRI's Network Analyst to identify the LCPs. This approach allows a more flexible cost model than ESRI's standard LCA procedures.

The LCPs created by Nolan and Cook (2012) show very clearly the drawbacks of applying LCA software (in this case: ArcGIS 9.3) that allows paths to move in eight directions only (like a queen on a chess board). On an isotropic cost surface with uniform costs, the LCP often deviates from the optimal straight line connection by moving in a cardinal direction (N, E, S, W) first and switching to a subcardinal direction (N-E, S-E, etc.) for the second part of the path (or vice versa). Many of the LCPs shown in this case study consist of long stretches in one cardinal or subcardinal direction; changing the orientation of the coordinate system by a few degrees could result in completely different paths (Herzog 2013b). The LCPs of Verhagen and Jeneson (2012) also show this effect.

Ullah and Bergin (2012) use the GRASS GIS procedure r.walk for calculating site catchments. This software optionally supports movements to 24 neighbours and therefore users can avoid the drawback mentioned above (but may skip a line-shaped high cost barrier unnoticed). In their case study, the cost of returning to the origin was not included and this is due to the fact that r.walk does not allow to adjust the cost function accordingly.

Livingood (2012) developed custom software for his cost model implementing both Dijkstra's algorithm and the A* algorithm. It is quite surprising that custom software was required to implement his fairly simple cost model. The Dijkstra algorithm identifies the globally optimal path between two locations (source and destination) in the graph if implemented correctly, although Kantner (2012) repeats the popular misconception that the algorithm favours the locally optimal solution. However, alternative LCPs accumulating the same costs or paths with insignificantly higher costs exist. These can be found by ESRI's corridor function applied by Surface-Evans (2012): the two cost surfaces for the source and the destination are added, corridors of low cost are identified by selecting the raster cells with the lowest 10% of this new raster.

ESRI software is based on Dijkstra's algorithm, GRASS r.walk supports both the locally optimal drain procedure and Dijkstra's method. It seems that IDRISI does not apply Dijkstra's algorithm and the anisotropic calculations assume that all downhill movements require less effort than travelling on flat terrain.

9. Varying parameters in the cost model

Refining the cost model often is required if the initial result does not fit with the validation data. Moreover, varying the parameters of the cost model allows checking the stability of the LCA results. Therefore, such additional calculations are important and should be included in a best practice standard for LCA studies (Herzog 2014a).

Rademaker, Reid, and Bromley (2012) experiment with several parameters (weight and load of the walker) of their cost function. They observe a «high degree of similarity» between these LCPs, which «compare favorably with those produced by the Tobler hiking function».

Verhagen and Jeneson (2012) present anisotropic and isotropic LCPs based on the Tobler function, moreover they combine this cost factor with a visibility component called openness.

Least-cost networks coinciding for some stretches but widely different at other locations are the result of Kantner's (2012) case study employing different grid sizes, Pandolf's formula, Tobler's hiking function and isotropic as well as anisotropic calculations. But no validation is given, so it is not possible to identify the best approach.

10. Validation

Often case studies do not include a validation of the LCA results (Table 1). Most authors presenting validations rely on the assumption that the location of known sites is related to past routes. An example is the case study of Nolan and Cook (2012): they analyse the agreement between the LCPs of their focal networks and known site locations and observe differences with respect to this agreement between the two focal settlements considered. Murrieta-Flores (2012) notes: «A high coincidence between settlements and corridors was an expected observation in general, as it is likely that human societies will locate their settlements within reach of communication networks».

In the case study of Phillips and Leckman (2012), ground truthing is performed by visually comparing the sherd distribution and the reconstructed paths. The find distribution maps create the impression that straight line connections outperform the reconstructed paths. Phillips and Leckman come to the conclusion that the physical landscape did not shape the paths, but that the social landscape probably was more important. However, a proper implementation of an appropriate cost function could have led to another conclusion.

Posluschny (2012) discusses the hypothesis that the location of Iron Age burial mounds is correlated with routes of that time. He shows that in his study area, most known burial mounds were recorded in forest areas and explains this observation by the ploughing activities in the agrarian area. So it is not always possible to validate reconstructed routes by comparing them to the distribution of selected indicator sites.

White (2012) records traces of old roads on satellite images and verifies the traces on the ground. This allows him to compare the results of his calculations with known route segments: he counts the segments of known routes aligned with reconstructed routes and calculates proportions. This is the most reliable way to check the LCA results, yet there is always some subjectivity in defining path segments and assessing alignment. A future task in LCA is to provide a more objective method for testing the agreement between a given network and the outcome of a reconstruction attempt.

11. Discussion and conclusion

Due to the space limitations several technical aspects could not be discussed in detail. For more details, the reader is referred to other publications (Herzog 2013a-b, 2014a-b).

The criticism of the case studies considered is on LCA methodology, although some of the studies apply LCA only as a tool in a larger methodological framework. In the latter case, the authors of the studies sometimes develop quite comprehensive models based on brilliant new ideas but neglect the LCA component. However, the weakest component in the model determines its overall performance.

A best-practice LCA case study applies reliable geographic data (representing the topography of the time frame considered) and adequate LCA software with an appropriate cost model. Often it is not possible to meet all of these requirements. Most important is the validation of the LCA results. As discussed above, route indicator sites like burial mounds provide only the second-best option, the best option is to compare the LCPs with remnants of old routes.

Irmela Herzog
The Rhineland Regional Council
The Rhineland Commission for Archaeological Monuments and Sites
Bonn

REFERENCES

Conolly J., Lake M. 2006, *Geographical Information Systems in Archaeology*, Cambridge, Cambridge University Press.

Earl G., Sly T., Chrysanthi A., Murrieta-Flores P., Papadopoulos C., Romanowska I., Wheatley D. (eds.) 2013, *Archaeology in the Digital Era. Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA) (Southampton 2012)*, Amsterdam, Amsterdam University Press.

Earl G., Sly T., Chrysanthi A., Murrieta-Flores P., Papadopoulos C., Romanowska I., Wheatley D. (eds.) 2014, *Archaeology in the Digital Era, Vol. II. e-Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA) (Southampton 2012)*, Amsterdam, Amsterdam University Press.

Fábrega Álvarez P., Parcero Oubiña C. 2007, *Proposals for an archaeological analysis of pathways and movement*, «Archeologia e Calcolatori», 18, 121-140.

Gaffney V., Stančič Z. 1992, *Diodorus Siculus and the island of Hvar, Dalmatia: Testing the text with GIS*, in G. Lock, J. Moffett (eds.), *Computer Applications and Quantitative Methods in Archaeology 1991*, BAR International Series 577, Oxford, Tempus Reparatum, 113-125.

Hare T.S. 2004, *Using measures of cost distance in the estimation of polity boundaries in the postclassic Yautepec Valley, Mexico*, «Journal of Archaeological Science», 31, 799-814.

Herzog I. 2013a, *Theory and practice of cost functions*, in F. Contreras, M. Farjas, F.J. Melero (eds.), *Fusion of Cultures. Proceedings of the 38th Annual Conference on Computer Applications and Quantitative Methods in Archaeology (Granada, Spain 2010)*, BAR International Series 2494, Oxford, Archaeopress, 375-382.

Herzog I. 2013b, *The potential and limits of optimal path analysis*, in A. Bevan, M. Lake (eds.), *Computational Approaches to Archaeological Spaces*, Walnut Creek, Left Coast Press, 179-211.

Herzog I. 2013c, *Least-cost networks*, in Earl *et al.* 2013, 240-251.

Herzog I. 2014a, *Least-cost Paths. Some Methodological Issues*, «Internet Archaeology», 36 (http://dx.doi.org/10.11141/ia.36.5).

Herzog I. 2014b, *Calculating accessibility*, in Earl *et al.* 2014 (online version), 720-734.

Hudson E.J. 2012, *Walking and watching*, in White, Surface-Evans 2012, 97-108.

Kantner J. 2012, *Realism, reality, and routes*, in White, Surface-Evans 2012, 225-238.

Kerig T. 2008, *Towards an econometrically informed archaeology: The Cologne Tableau (KöTa)*, in A. Posluschny, K. Lambers, I. Herzog (eds.), *Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA) (Berlin 2007),* Kolloquien zur Vor- und Frühgeschichte 10, 372 (full paper on CD).

Kienlin T.L., Cappenberg K., Korczynska M.M. 2012, *Überlegungen zu den spätbronze- und früheisenzeitlichen Landnutzungsstrategien im mittleren Dunajectal, Kleinpolen*, Festschrift für Hans Lohmann zugeeignet von seinen Schülern, Freunden und Kollegen zu seinem 65. Geburtstag (Paderborn 2012).

Livingood P. 2012, *No crows made mounds*, in White, Surface-Evans 2012, 174-187.

Murrieta-Flores P. 2012, *Understanding human movement through spatial technologies. The role of natural areas of transit in the Late Prehistory of South-western Iberia*, «Trabajos de Prehistoria», 69, 1, 103-122.

Nolan K.C., Cook R.A. 2012, *A method for multiple cost surface evaluation of a model of fort ancient interaction*, in White, Surface-Evans 2012, 67-93.

Phillips S.M., Leckman P.O. 2012, *Wandering the desert*, in White, Surface-Evans 2012, 46-66.

Posluschny A. 2012, *Von Nah und Fern? Methodische Aspekte zur Wegeforschung*, in O. Dally, F. Fless, R. Haensch, F. Pirson, S. Sievers (eds.), *Politische Räume in vormodernen Gesellschaften. Gestaltung - Wahrnehmung - Funktion. Internationale Tagung des DAI und des DFG-Exzellenzclusters TOPOI (Berlin 2009),* Rahden/Westf., Verlag Marie Leidorf, 113-124

Rademaker K., Reid D.A., Bromley G.R.M. 2012, *Connecting the dots*, in White, Surface-Evans 2012, 32-45.

Richards-Rissetto H. 2012, *Social interaction at the Maya site of Copán, Honduras*, in White, Surface-Evans 2012, 109-127.

Rissetto J.D. 2012, *Using Least Cost Path Analysis to reinterpret later Upper Paleolithic hunter-gatherer procurement zones in Northern Spain*, in White, Surface-Evans 2012, 11-31.

Smith M.J. de, Goodchild M.F., Longley P.A. 2007, *Geospatial Analysis. A Comprehensive Guide to Principles, Techniques and Software Tools*, Leicester, Matador.

Surface-Evans S. 2012, *Cost catchments*, in White, Surface-Evans 2012, 128-151.

Ullah I.I., Bergin S.M. 2012, *Modeling the consequences of village site location*, in White, Surface-Evans 2012, 155-173.

Verhagen P., Jeneson K. 2012, *A Roman puzzle. Trying to find the Via Belgica with GIS*, in A. Chrysanthi, P. Murrieta-Flores, C. Papadopoulos (eds.), *Thinking Beyond the Tool*, BAR International Series 2344, Oxford, Archaeopress, 123-130.

White D.A. 2012, *Prehistoric trail networks of the Western Papaguería*, in White, Surface-Evans 2012, 188-206.

White D.A., Surface-Evans S. (eds.) 2012, *Least Cost Analysis of Social Landscapes. Archaeological Case Studies*, Salt Lake City, The University of Utah Press.

ABSTRACT

The application of least-cost analysis (LCA) in archaeology has considerably increased in recent years. Modern Geographical Information Systems provide the tools for generating least-cost site catchments, least-cost paths and route networks as well as accessibility maps. Recently, published case studies present LCA results for very different time periods and parts of the world. Consequently, it seems that the technology for generating these results is readily available and reliable. However, the quality of the LCA outcome depends on the accuracy and the resolution of the geographical data used, and on the cost model itself. Varying the parameters of the cost model allows assessing the stability of the modelled catchments, routes or accessibility maps. Without validation, the LCA results remain exploratory and should not be used as a basis for building an even more complex model. The technical aspects of the case studies considered will be discussed with respect to these issues.