Tilt – Slope-Dependent Least Cost Path Calculations Revisited

Irmela HERZOG[1] – Axel POSLUSCHNY[2]

[1]The Rhineland Regional Council,
The Rhineland Commission for Archaeological Monuments and Sites, Bonn, Germany
[2]Roman-Germanic Commission, Frankfurt, Germany
[1]i.herzog@LVR.de
[2]posluschny@rgk.dainst.de

Abstract

Least-cost path (LCP) calculations have become a standard procedure in archaeology and this is why many publications simply mention the name of the GIS software package used in the analysis they perform but do not give details of the parameters selected. Most LCP calculations rely heavily on slope values derived from a DEM of modern landscape. However, the LCPs generated this way will vary depending on many factors including DEM resolution and accuracy, slope calculation method, slope-dependent cost function, method for dealing with anisotropy, and the number of neighbours considered in the grid. This paper gives an overview of problems encountered in slope-dependent LCP calculations using the Wetterau region in Germany as a study area.

Keywords

Least-cost paths, DEM, slope

1. Introduction

Least-cost path (LCP) calculations have become a standard procedure in archaeological GIS studies. The aim is to reconstruct the location of ancient roads and pathways when only the end and start points, and sometimes part of the ancient path, are preserved (e.g. van Leusen 2002, 16–10 to 16–18). Alternatively, optimal paths are calculated to understand the principles of path formation and to compare the paths to well-documented routes used in the past (e.g. Bell and Lock 2000).

As Rahn (2005) points out, the typical GIS user can now pick and choose between various equations for optimal path calculation and implement them directly using a variety of GIS packages (Arc/INFO-GIS, GRASS-GIS, Idrisi, etc.), without this task requiring much in the way of specialist knowledge or training. Writing a specialised program or script to accomplish this is no longer necessary because the functionality has already been built into the package as purchased. The other side of the coin is described by Branting (2004, 27): "But increasingly today it is possible to click a button and have the analysis done for us, whether or not we have any idea what we are doing or what the results mean." We claim that some expertise is required to set up the equations, to become aware of the limits of the GIS procedures applied, and to analyse the results.

While we assume that in prehistory people tried to optimise the costs of routes which were taken frequently, we cannot be sure that people actually found the LCP; they might rather have been satisfied with a path involving slightly higher costs – a problem that is not addressed in this paper.

Calculating the LCP between two given locations is easier than generating a LCP network that interconnects a larger number of locations. We will focus on the simpler task since the algorithm for solving this problem is well-known and was presented by Dijkstra nearly fifty years ago (Cormen *et al.* 2001, 595–599). The network problem can be solved by connecting every location with every other location in the network by a LCP, but this means that the crisscrossing of paths will occupy lots of space, which might not be desirable in an agrarian landscape. An alternative is to minimise the total cost of the network of paths connecting every location with every other location, which was proposed by Jens Andresen at his CAA 2008 talk.

We analysed quite a few LCP studies, and of those which give details of the cost factors involved take into account the slope of the terrain, the only exception is an optimal path analysis in an extremely flat landscape. Therefore, we concentrate our efforts on slope-dependent LCPs, though we are aware of the fact that other environmental factors like the presence of streams, the type of terrain (sand, bog, vegetation) and social factors like taboo zones or attractors or existing transportation facilities play an important role as well. Most of the publications on LCPs in archaeology calculate routes in a rural area; only recently have studies of urban pedestrian movements been presented (e.g. Branting 2004). Because in a city the direction of movement is limited to the empty spaces between the buildings and walls, the optimal route algorithm has to be set up differently than in rural areas, on which this paper focuses.

As practical examples, we calculate LCPs connecting four Early Latène sites in the vicinity of the famous proto-celtic hillfort Glauberg in the Wetterau, Hesse, Germany (Posluschny 2007; *Fig. 1*). The DEM has a resolution of 25m and was supplied by the German Federal Agency for Cartography and Geodesy (www.bkg.bund.de). The accuracy of the elevation values is in the range of ±1 to 5m.

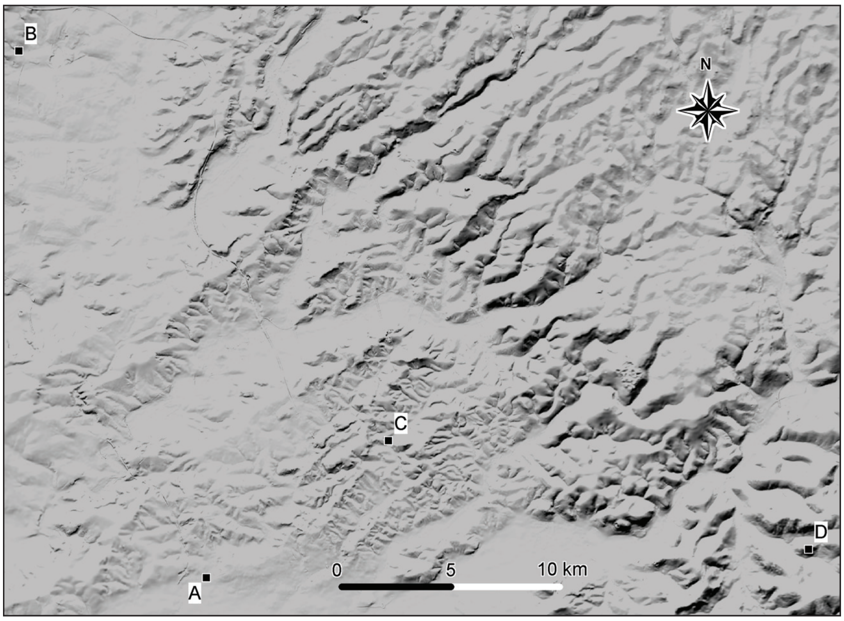

Fig. 1. The test area with four Latène sites. The elevation of A is 113m asl., B 176m asl., C 164m asl., and D 298m asl.

2. Dijkstra's algorithm

Dijkstra's algorithm requires the cost of every subpath to be non-negative, which is obviously the case in archaeological LCP studies, because every move requires time and energy. Dijkstra's algorithm has been implemented by Llobera (2000) for his model describing the sociology of movement, and was available in ESRI's ArcInfo software as early as in 2000 (Harris 2000). However, other GIS software packages apply different procedures which are not guaranteed to provide the correct shortest path. GRASS GIS procedures r.walk (http://www.grass.itc.it/grass63/manuals/html63_user/r.walk.html) and r.cost (http://www.grass.itc.it/grass63/manuals/html63_user/r.cost.html), for example, use only part of Dijkstra's algorithm, combining it with a drainage procedure which might get caught in local minima. It is for this reason that Conolly and Lake (2006, 254) note that LCPs often do not appear to follow the globally optimal route. But if Dijkstra's algorithm is implemented properly the optimal path will be found.

Dijkstra's algorithm was designed for networks of paths where the cost of traversing each subpath is known. The cost of each path is the sum total of the subpath costs. Archaeological LCP calculations are normally based on raster data, and the raster data is transformed to vector data by connecting each raster cell with its neighbouring cells by a (virtual) vector or subpath. This transformation introduces an error, as the accuracy of the LCP calculation depends not only on the correct implementation of Dijkstra's algorithm, but also on the number of neighbours considered for each cell. The larger the number of neighbours the smaller the elongation error, i.e. the worst case difference in length between the LCP and the optimal route. According to Huber and Church (1985) the elongation distortion for an eight-neighbour network may exceed 8% of the optimal route length, whereas with 24 neighbours, the worst case is 2.8%. When 48 neighbours are considered, the elongation error is below 1.4%. ESRI's ArcInfo software supports only eight neighbours, whereas the GRASS GIS procedures r.walk and r.cost optionally work with 24 neighbours.

Fig. 2 shows an example of LCPs calculated with different numbers of neighbours in our study area. Whereas the 24 and 48-neighbour paths follow the same route most of the time, the path from A to B with eight neighbours deviates significantly from the other two routes. The length of the eight-neighbour path between A and B is 27.8km, whereas the 24-neighbour path has a total length of 26.3km, and the 48-neighbour path is only somewhat shorter (26.2km). For the eight-neighbour path between A and B, the total costs exceed by 4.8% those of the

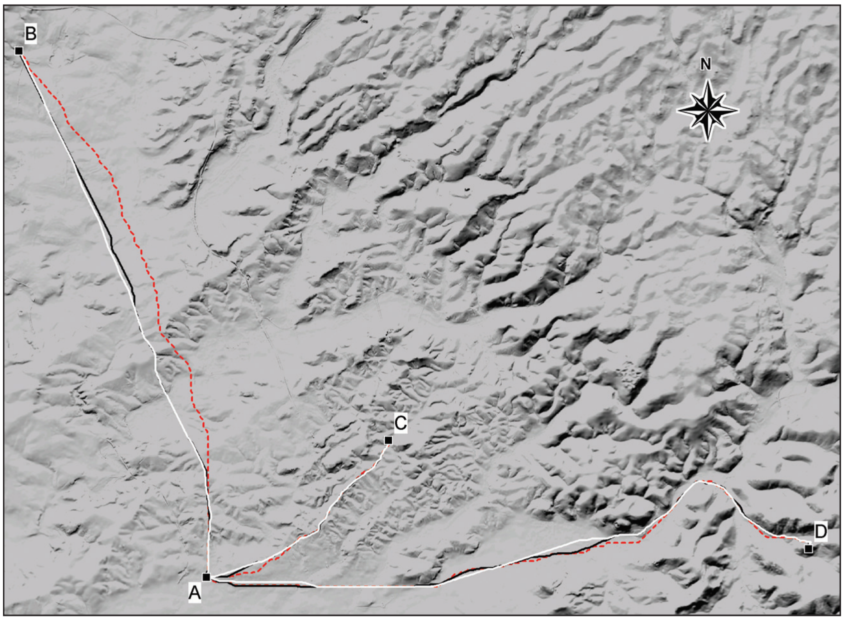

Fig. 2. Comparison of LCPs calculated with taking eight (dotted), 24 (black) and 48 (white) neighbours into account.

resolution DEM comes closer to the human perception of a landscape (a subject worthy of further research).

A method to assess the error introduced by the low resolution grid is shown in *Fig. 3* which compares the histogram of slope values for the 25m-grid and a 100m-grid of the study area, where the elevation points of the 100m-grid form a subset of the 25m-grid points. With the 25m-grid, slopes steeper than 20% can be found in 8.8% of the study area, compared to 7.1% when taking the slope map of the 100m-grid as a basis for this calculation. It appears plausible that a lower resolution grid always results in some smoothing of the DEM.

48-neighbour path. Similarly, the eight-neighbour LCP from A to D expends 4.3% more costs than the 48-neighbour path. On increasing the number of neighbours considered the computation time increases as well: With our program, which is still in the development stage, the computation time of the 24-neighbour paths was 1.65 times as long as that of the eight-neighbour paths, and with 48 neighbours, this factor became 2.6.

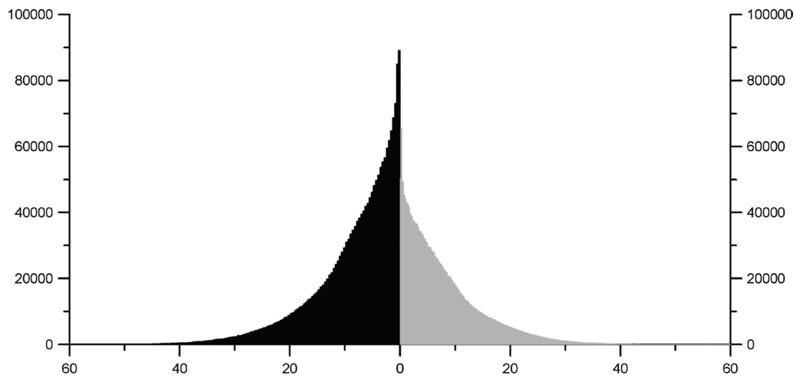

Fig. 3. Comparison of the slope distribution of the 25m-grid (left) and the 100m-grid (right) for the study area shown in Fig. 1.

3. The DEM – basis for calculating slope values

The calculation of slope depends on the accuracy and resolution of the DEM, and the problems of DEM generation are well-known (Beex 2004; Wheatley and Gillings 2002, 158), but only few authors presenting DEM related optimal path studies address this issue. An exception is Branting (2004, 61–4), who puts much effort into setting up a DEM with an error well below the level of the individual step length at ca. 60cm. Assuming that every step of a pedestrian (with a step length of 60cm) is to be modelled and applying the Nyquist limit as proposed by Beex (2004), a DEM with a minimum distance of 15cm between elevation points is required. Of course, DEMs with such a high resolution are only available in rare cases, and with a research area of 100km² or more, the storage requirements and computational load will become prohibitive. Some archaeologists believe that a low

When comparing the slope values of the 25m-grid and the 100m-grid, the slope difference exceeds 1% within 36% of the test area. For 19% of the test area, the difference is greater than 2%, and a difference of more than 5% can be found in 4% of the study area. The consequences of using various cost surface resolutions have already been discussed by Huber and Church (1985). Harris (2000) reports that one of his experiments with LCP calculations based on DEMs with diverse resolutions produced two very different LCP corridors.

We performed a similar experiment, calculating LCPs on the basis of the 25m- and the 100m-grid (*Fig. 4*). In two cases, the LCPs produced by different cell sizes are similar, but the path connecting A and C follows the ridge for the 25m-grid and is located in the valley for the 100m-grid. The LCPs of the low resolution grid are shorter than the high resolution

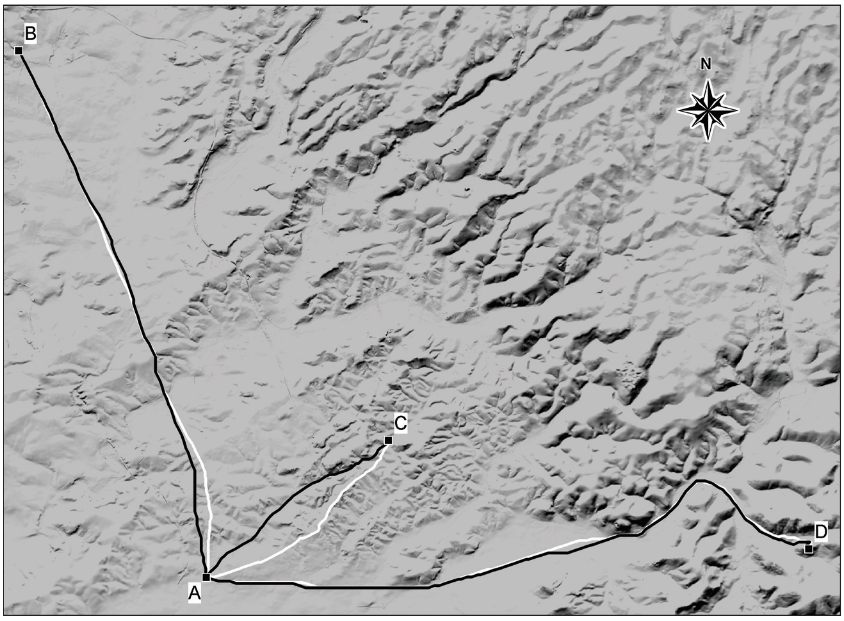

Fig. 4. Comparison of LCPs calculated on the basis of the 25m DEM (white) and the 100m DEM (black).

paths. For example, the length of the path from A to C is 10.8km for the 100m-grid but 11.1km for the 25m-grid.

Another way to assess the impact of the DEM errors is to experiment with small modifications of the DEM. A random error within the limits of the DEM accuracy might be added to the elevations and the LCP calculation repeated with this slightly modified DEM. In addition, minor changes of the DEM might be introduced by resampling. The stability of the LCP calculation is estimated by comparing the initial LCP with the optimal routes produced by the modified DEMs. *Fig. 5* shows the outcome of one of our experiments with a DEM modified by a random error of ±2m. The LCPs are similar to those presented in *Fig. 4*: Whereas the LCPs connecting A with B and A with D are fairly stable, the modified DEM results in an alternative route from A to C.

In addition, the relief of an ancient landscape often underwent substantial changes compared to the modern landscape, owing to both natural forces and human activities. Even during the Neolithic period agrarian use of the landscape entailed significant erosion. Sometimes, archaeological sites are well preserved under a layer of 5m of eroded soil from somewhere else (colluvium) and are only discovered by exceptional deep excavations (e.g. in connection with underground car parks or open-cast mining). This example shows that many archaeologists actually have seen a relief change of 5m or more. Moreover, quarrying and small-scale open-cast mining, pits dug for tile clay or calcareous soils, modern road building and many other human activities have left substantial traces in modern relief. When taking change by natural and cultural factors into account, reconstructing the DEM of a certain period becomes a very difficult and time-consuming task (Gerlach 2003). In the DEM we are using in this study, man-made structures can be seen, e.g. the route of motorway 45 starting in the north-western corner of the test area and moving in long curves to the ridge between sites A and B, after that following a straight course to the ridge connecting sites A and C. In addition, some larger pits dug for the extraction of bulk material can be identified easily, for example 11.5km to the east of site C on a hilltop. However, reconstructing the Iron Age landscape for the test area is beyond the scope of this paper.

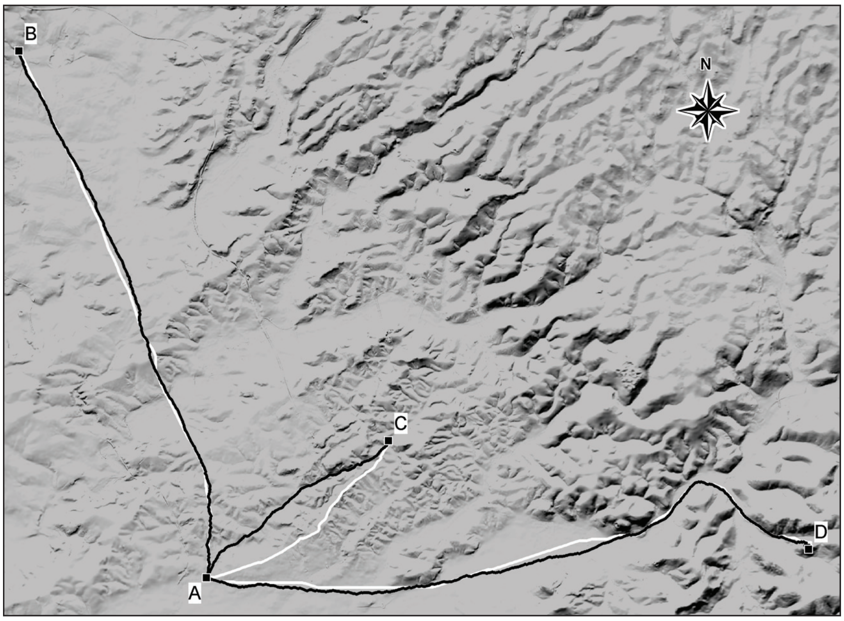

Fig. 5. Comparison of LCPs calculated on the basis of the 25m DEM (white) and the DEM modified by a random error of ±2m (black).

4. Slope calculation issues

Kvamme (1992, 129) notes that different algorithms exist to calculate slope. But in the archaeological LCP papers we have studied, the algorithms chosen for slope estimation and the quality of the results are not discussed. For example, the slope value calculated in ArcView is the maximum rate of change between each cell and its neighbours, whereas the slope map of Vertical Mapper, an add-on for the GIS software MapInfo, shows the average slope between each cell and its neighbours. Warren et al. (2004) compared the slope data calculated by ten experts using their preferred GIS tools with the true values measured in the field. The 143 ha study area was a hilly landscape with slopes ranging up to 30%; 4500 elevation points form the basis for the DEMs. The experts' task was to estimate the slope of 57 points randomly distributed across the central part of the study area. The square root of the mean of squared differences between estimated and measured slopes for the sample locations quantifies the error of each GIS tool. The mean error value was in the range of 2 to 3%, and the authors of the study come to the conclusion that variation in the computation of slope from digital elevation data can result in significantly differing slope values.

Major errors are introduced if the units for measuring slope are confused. Mathematical slope is commonly taught as "rise over run" or rise/run; this value multiplied by 100 results in percent slope. Mathematical slope is the tangent of angle slope, and the degree and the radian are the most common units used to measure angles.

5. Cost functions

One of the most popular cost curves used in archaeological LCP studies is the Tobler hiking function which was first introduced to archaeologists by Gorenflo and Gale (1990). The formula calculates the walking speed depending on slope, and from this, the time requirements can be easily determined. In the case of the Tobler hiking function confusion of units for measuring slope is wide-spread: In the original publication slope is calculated by vertical change divided by horizontal change, i.e. mathematical slope, but the unmodified formula is cited by van Leusen (2002, 6–6) as well as Connelly and Lake (2006, 219) claiming that slope is measured in degrees, whereas Wheatley and Gillings (2002, 155) present the Tobler formula and refer to percent slope.

Anyway, for modelling walking costs, we prefer to employ a function that is based on energy rather than time expenditure. In our view, the formula presented by Llobera and Sluckin (2007) which is based on a large sample of metabolic cost measurements published by Margaria in 1938, is the most reliable cost function for pedestrian movement currently available. We assume that the same route was taken on both ways, from A to B and from B to A.

Whereas pack animals and pedestrians can be modelled with the same cost curve, this cost curve is not appropriate if carts or wagons are drawn on the routes. Wheeled traffic cannot climb steep slopes as easily as pedestrians do. Renfrew and Bahn (1996, 315) note that wheeled vehicles first appeared in the 4th millennium BC in the area between the Rhine and the Tigris. Wheels were made of wood at that time and only very few wooden objects were preserved due to exceptional conditions. For this reason, it is difficult to decide whether the Latène people considered in our example used vehicles for transport on a regular basis. Descriptions of historic routes for wheeled transport often include an estimate of the critical slope (i.e. the transition where switchbacks become more effective than direct uphill or downhill paths), and hardly any other information is available. This estimate is often in the range of 8 to 12%, and therefore we chose a critical slope of 10% for the examples presented in *Figs 2*, *4* and *5*. The cost curve for a given critical

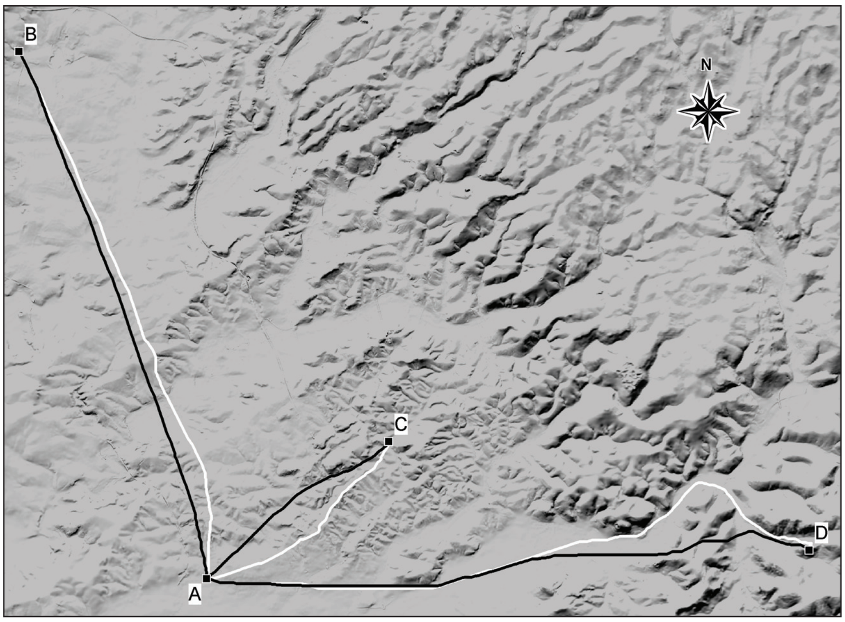

Fig. 6: Pedestrian LCPs (black) compared to vehicle LCPs with a critical slope of 10% (white).

slope can be constructed based on the formulas given by Llobera and Sluckin (2007). Compared to roads for vehicles, the routes for pedestrians are usually more direct (*Fig. 6*).

6. Conclusions

Different cost functions are appropriate for modelling vehicle and pedestrian movement. Several methods are available to test the stability of LCP calculation results. The examples presented in *Fig. 5* show that a very different LCP may result after modifying the DEM within the limits of its accuracy. Both popular GIS software packages, ArcView and GRASS, have significant drawbacks when it comes to LCP procedures: whereas ArcView results can be distorted by substantial elongation errors due to the small number of neighbours considered, in GRASS Dijkstra's algorithm is implemented incompletely so that the computed path may involve more costs than the true LCP. This paper presents results from an ongoing project which will soon be published in more detail.

Acknowledgements

We would like to thank Karsten Lambers, Martin Sauerbier, Benjamin Ducke, Jens Andresen, and Ingo Eichfeld for providing us with useful references and for helpful suggestions.

References

Beex, Willem (2004). Use and Abuse of Digital Terrain/Elevation Models. In: Magistrat Stadt Wien (eds.) *Enter the Past. The E-way into the Four Dimensions of Cultural Heritage. Proceedings of the 31st CAA Conference, Vienna, 2003* (BAR International Series 1227) Oxford: ArchaeoPress, 240–242.

Bell, Tyler and Gary Lock (2000). Topographic and cultural influences on walking the Ridgeway in later prehistoric times. In: Gary Lock (ed.) *Beyond the Map: Archaeology and Spatial Technologies*. (NATO Science Series, Series A 321.) Amsterdam: IOS Press, 85–100.

Branting, Scott Andrew (2004). *Iron Age pedestrians at Kerkenes Dag: An archaeological GIS-T approach to movement and transportation (Turkey, England)*. PhD Thesis: The State University of New York at Buffalo.

Conolly, James and Mark Lake (2006). *Geographical Information Systems in Archaeology*. (Cambridge Manuals in Archaeology) Cambridge: Cambridge University Press.

Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest and Clifford Stein (2001). *Introduction to Algorithms*. Second Edition. Cambridge: The MIT Press, McGraw-Hill Book Company.

Gerlach, Renate (2003). Wie dynamisch sind die geogenen Grundlagen einer archäologischen Prognose? Die Veränderungen von Relief, Boden und Wasser seit dem Neolithikum. In: Jürgen Kunow and Johannes Müller (eds.) *Archäoprognose Brandenburg I*. (Forschungen zur Archäologie im Land Brandenburg 8) Wünsdorf, 89–96.

Gorenflo, Larry J. and Nathan Gale (1990). Mapping Regional Settlement in Information Space. *Journal of Anthropological Archaeology* 9, 240–274.

Harris, Trevor (2000). Moving GIS: exploring movements within prehistoric cultural landscapes using GIS. In: Gary Lock (ed.) *Beyond the Map: Archaeology and Spatial Technologies*. (NATO Science Series, Series A 321.) Amsterdam: IOS Press, 116–123.

Huber, Dennis L. and Richard L. Church (1985). Transmission corridor location modeling. *Journal of Transportation Engineering* 111 (2), 114–130.

Kvamme, Kenneth L. (1992). Terrain Form Analysis of archaeological location through Geographic Information Systems. In: Gary Lock and Jonathan Moffet (eds.) *CAA 91: Computer Applications and Quantitative Methods in Archaeology 1991* (BAR International Series S577) 127–136.

Leusen, Martijn van (2002). *Pattern to Process. Methodological Investigations into the Formation and Interpretation of Spatial Patterns in Archaeological Landscapes*. PhD Thesis: Rijksuniversiteit Groningen (http://irs.ub.rug.nl/ppn/239009177)

Llobera, Marcos (2000). Understanding movement: a pilot model towards the sociology of movement. In: Gary Lock (ed.) *Beyond the Map: Archaeology and Spatial Technologies*. (NATO Science Series, Series A 321.) Amsterdam: IOS Press, 65–84.

Llobera, Marcos and T. J. Sluckin (2007). Zigzagging: Theoretical insights on climbing strategies. *Journal of Theoretical Biology* 249, 206–217.

Posluschny, Axel (2007). From Landscape Archaeology to Social Archaeology. Finding patterns to explain the development of Early Celtic „Princely Sites“ in Middle Europe. In: Jeffrey T. Clark and Emily Hagemeister (eds.) *Digital Discovery. Exploring New Frontiers in Human Heritage. CAA 2006 Computer Applications and Quantitative Methods in Archaeology.* Proceedings of the 34th Conference, Fargo, United States, April 2006. Budapest: Archaeolingua, 131–141.

Rahn, Brian (2005). Cost Surface-Derived Least-Cost Paths: A Case Study from Iron Age Orkney. *Internet Archaeology* (http://intarch.ac.uk/journal/issue19/rahn_index.html)

Renfrew, Colin and Paul Bahn (1996). *Archaeology. Theories, Methods, and Practice.* London: Thames and Hudson Ltd.

Warren, S. D., M. G. Hohmann, K. Auerswald and H. Mitasova (2004). An evaluation of methods to determine slope using digital elevation data. *Catena* 58, 215–233.

Wheatley, David W. and Mark Gillings (2002). *Spatial Technology and Archaeology: The Archaeological Applications of GIS.* London: Taylor & Francis.