

## Appendix B: Least-cost path (LCP) calculations

In this section we present a more detailed background of the LCP model we have developed for this paper. The flowchart (Fig. B1) shows the workflow applied for generating the LCPs (Section 4.3.).

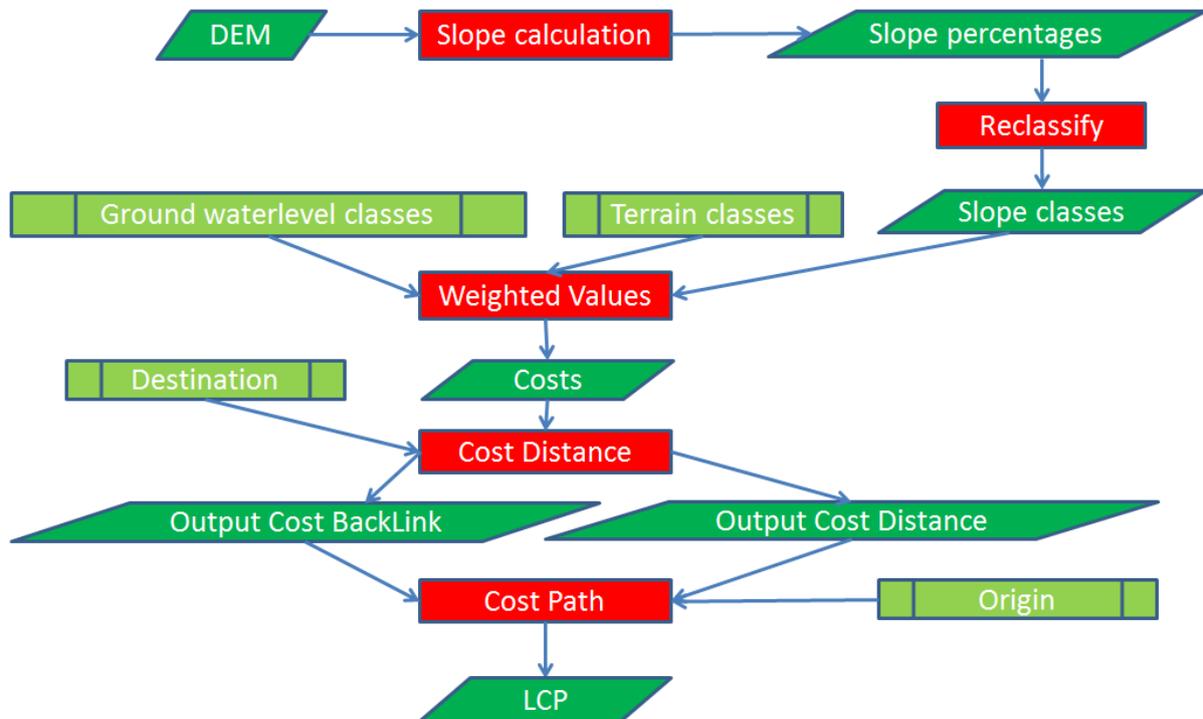


Figure 1: Flowchart of the LCP model

As mentioned before (Section 4.3.), there are three factors influencing the outcome of the model: terrain, slope and groundwater levels. The factors were classified prior to the calculation of the combined friction raster. It is obvious that the allocations of the values to these classes have a crucial impact on the outcome of the model.

Slope values were derived from the digital terrain model (DTM) of the study area as percentage units of measurement. Subsequently, slope percentages were grouped by 'Natural breaks' classification, a statistical method developed to establish the optimal arrangement of values into different classes (Jenks, 1967). Classes representing a group of higher slope percentages were allocated with a higher value (cost) than the one representing lower percentages. Slope percentages higher than 10% were excluded from calculation as the roads of that time probably did not include switchbacks allowing to climb steeper slopes. This number is based on the average of the range of 8% and 12% threshold in Herzog and Posluschny (2012). In the end it resulted in 10 classes. The class with the highest slope had a cost of 10, the second highest a cost of 9, etc.

The terrain classes needed refining before they were incorporated in the model. Terrain factor is a combination of topographic depressions and lower elevation areas, both of which could represent wetter areas. Depressions were identified with the 'fill' and 'cut fill' tools in ArcGIS. Using the fill tool, depressions based on the DTM were filled to provide a smoother surface. Subsequently, the fillings (i.e. the depressions) were calculated by subtracting the original DTM from the filled DTM. A threshold factor of 6 meters above sea level was employed to distinguish wetter areas from (drier areas and is based on a combination of the soil map, historical 1850 map and the DTM data.

*The threshold was chosen by overlaying these sources and determining the height of the boundary between sand and clay areas; the transition zone from the push moraine to the coastal zone was taken into particular account. The raster calculator tool then divided the DTM in areas below and above the threshold factor. The result was combined with the depressions map and produced a terrain map refined into lower and higher areas. The costs of these two classes are based on the terrain coefficients from Soule and Goldman (1972): 1.2 for the higher sandy heath land and 1.8 for lower wetlands.*

*The groundwater level classes were based on the groundwater-level table of the soil map and needed a small adjustment in ArcGIS. In other words, when a class represents a deep groundwater level a lower value (cost) was allocated. For this purpose the 'reclassify' tool was used.*

*The groundwater level exist nine classes including subclasses. To make them suitable for the model these has to be regrouped to seven classes, where the lowest groundwater class represents the lowest cost of 1, the second lowest a cost of 2 and so on.*

*The weighted values define the final costs and a least cost path was calculated with the standard cost distance tools for LCP in ArcGIS between and origin and destination. Naturally, the division of weighted values has an enormous influence on the outcome. The optimal value for each route was achieved by a simple 'trial and error' exercise. This is the most laborious and time consuming part of the model. Nevertheless, this flexibility and the overall flexibility of the model to change values are an asset. Indeed, it is quite easy to investigate a certain factor or make a comparison between factors for a chosen route. Although one should take into account that it is not only the set parameters that define the outcome, but also the software package used (Gietl et al., 2008). The model can be run on a standard desktop computer, which is advantage over other models which require more processing power (Verhagen, 2012; White and Barber, 2012).*

*In order to make a good evaluation of the LCP model, four historical roads were reconstructed: the Hessenweg between Hattem and Voorthuizen, which connected Amsterdam overland to Germany, the road (Harderwijkerweg) between Harderwijk en Arhem and two routes across the Veluwe; from Dieren to Barneveld, and from Apeldoorn to Voorhuizen. These roads have been chosen as they have known historical depth and cross the Veluwe push moraine in different directions and in different routes: from Northeast to Southwest over the brink of the push moraine, from North to South along the western border and from East to West across the central part. For each of these routes different parameter sets were tested. In additon to the 100m route-zone width presented in the main text, we also applied a 200m and 1000m route zone to investigate the four routes further and to make a broader comparison.*

**Table 1: Correlation per LCP-calculated routes (varying route-zone widths), the ALS-extracted hollow ways, the AD 1600 route network and the TMK 1850. When a percentage is in underlined, it indicates that a route performed best in comparison to the others.**

Dataset	Route: Arnhem to Harderwijk		
	100m zone	200m zone	1000m zone
ALS-extracted hollow ways	7 %	<u>14%</u>	26%
AD 1600 route network	9%	20%	<u>67%</u>
TMK 1850 routes	<u>58%</u>	<u>72%</u>	<u>88%</u>

Dataset	Route: Hattem to Voorthuizen		
	100m zone	200m zone	1000m zone
ALS-extracted hollow ways	<u>13%</u>	<u>14%</u>	28%
AD 1600 route network	10%	25%	46%
TMK 1850 routes	30%	50%	<u>89%</u>

Dataset	Route: Apeldoorn to Voorthuizen		
	100m zone	200m zone	1000m zone
ALS-extracted hollow ways	9%	13%	<u>39%</u>
AD 1600 route network	<u>21%</u>	<u>30%</u>	50%
TMK 1850 routes	32%	46%	79%

Dataset	Route: Dieren to Barneveld		
	100m zone	200m zone	1000m zone
ALS-extracted hollow ways	1%	1%	1%
AD 1600 route network	3%	7%	47%
TMK 1850 routes	11%	21%	70%

**Table 2 provides an overview of correlation percentages per LCP-calculated route. Although the general correlation is highest with the TMK 1850 dataset and the lowest with the ALS-extracted hollow ways, the individual LCP routes show varying results per different route-zone width (buffer). This is not surprising because the 1850 map shows a much higher number of (alternative) roads and paths and the LCP routes we selected were also known to be present in the landscape during that time.**

**In general we can state that the reconstructed route Arnhem to Harderwijk performed best, followed up by the Hattem to Voorthuizen modelling. However, the modelled Apeldoorn to Voorthuizen route performed best regarding the AD 1600 network. A premature interpretation may suggest that the Apeldoorn Voorthuizen route is older than others, this however is not substantiated by any further evidence and is beyond the limits of these datasets to provide an answer.**

**Table 2: The weighted values of the factors per route. Where a percentage is underlined, it indicates this factor was most influential for a route.**

Routes	Factors		
	Slope	Terrain (low areas)	Ground Water level
Arnhem to Harderwijk	20%	40%	40%
Hattem to Voorthuizen	5%	90%	5%
Apeldoorn to Voorthuizen	5%	90%	5%
Dieren to Barneveld	33%	33%	33%

*It is clear that the terrain, representing the low areas, is the most influential factor followed up by the ground water level in this model. However, it is also evident that slope has certain influence in all the routes. The table also demonstrates clearly there is no fixed division of values for routes across the Veluwe. Indeed, it advocates the individual investigation of routes. This may also indicate that the creation of a route network for an area is per definition limited, as individual routes may have their division of factors of influence.*

*The individual LCP routes are shortly described below.*

**Route: Harderwijk to Arnhem**

*If we look at the LCP reconstruction of the Harderwijkerweg, it is clear that in the beginning, from Arnhem, it follows the old Harderwijkerweg better than the more recent route on the TMK 1850 map. These old and new roads come together at a certain point and the modelled route stays reasonably close by. However, after a while there is a divergence, which can be explained by the lower areas class of the terrain factor, which the reconstructed path avoids and the real road runs over it. In the final part the divergence is largest. This may be due to cultural factors, meaning that in the real route a village was included, which led to a small detour.*

*The overlap with the 1600 network is clearly minor. As mentioned before this is mainly due to lesser density of this network. Increasing the ....zone improves the overlap. The same accounts for the overlap with ALS extracted roads.*

**Route: Hattem to Voorthuizen**

*The LCP itself runs in the beginning between the old and new Hessenweg on the TMK map and then it keeps a certain distance to it and in the final part there is quite a distance between the original route and the path. This has perhaps to do with the fact that the reconstructed path is quite straight and the original makes a little detour to the village of Garderen. This may be due the influence of cultural factors.*

*Regarding the overlap with the 1600 the overlap is doubled with a zone of 200 meter is used. A further increase results in a relative small improvement. It is the best route lining up with the extracted roads with a 100m zone. However, this picture changes when larger...zones are used.*

**Route: Apeldoorn to Voorthuizen**

*The third reconstructed road was the one from Apeldoorn to Voorthuizen. This one shows already within itself a lot of variation when confronted with the comparison data sets. Collating to the 1600 map shows that the beginning of the LCP does not come close to a road. In the central part, it is approximately parallel before losing track again in the final part. The comparison with the 1850 shows a different picture. There was generally more overlap. Therefore, the reconstruction for this map was quite good, although still not so good as the one of the Harderwijkerweg.*

*Regarding the 1600 network, this route performed best for the smallest route zone of 100 meters. This could indicate something over the age of this route. Also this route lines up quite well with the extracted roads when larger route zones are used.*

**Route: Dieren to Barneveld**

*The reconstruction of the Dieren Barneveld is the least successful of all least cost paths. Especially at the start the modelled route runs around the Caloumen Berg in a northern route, but the map of 1600 shows a southern turn, while comparison to the 1850 map shows little overlap with a road that could be depicted as regional. In the final part, its performance improves and it follows the corridor between the sand blown areas on the 1850 map quite well. Overlap with hollow ways however, is almost not existent, indicating perhaps that they are less represented in this area, or they have simply vanished with time.*

#### *Reference*

*Jenks, G. 1967. The Data Model Concept in Statistical Mapping. In International Yearbook of Cartography 7: 186–190.*