Modeling slope as a contributor to route selection in mountainous areas

Thomas J. Pingel
Department of Geography
University of California Santa Barbara
Phone: (805) 893-3663
Fax: (805) 893-3146
E-mail: pingel@geog.ucsb.edu

Abstract:
Slope exerts a powerful influence on the route selection processes of humans. Attempts to model human movement in hilly and mountainous terrain that have largely focused on least-time route transformations can be improved by incorporating research that suggests humans systematically overestimate slopes. Such research suggests that cost functions derived from slope should be more expensive than time derivations alone would indicate. This paper presents a method that empirically estimates cost functions for slopes. The method is then used to predict routes and paths that are more likely to be selected by humans based on their perceptions of slope. We also evaluate that method and find it successfully predicts road, track and trail locations over a variety of conditions and distances.

1. Introduction

Route selection is one of the most basic and ubiquitous problems faced by humans, yet one of the primary influences on long-distance travel - overcoming slope - is not well considered. When people list important criteria in route selection, slope is typically either omitted (Golledge 1999) or deemphasized (Hochmair 2004) as a significant contributor. The reason for this is probably that most human travel occurs within familiar, confined, and typically low-relief areas. Further, most human travel in these spaces occurs on networks of paths, which are themselves sited away from high slope areas.

What we consider here is the contribution of slope in a continuous and direction-sensitive setting with no particular regard for preexisting networks. This allows the considerations of problems related to autonomous navigation by robots and the reasons for the development of paths by prehistorical civilizations and animals. While we consider human route selection explicitly, many of the principles presented here can apply to any agent.

One way to describe how humans choose to move through their environment is to envision the medium of travel as filled with costs of different types with agents choosing routes that minimize the total cost. Distance is the most fundamental cost of moving through a space, but humans select routes based on more than just distance. One other contributing criterion is the slope of the surface since flatter terrain allows for more direct, faster, and easier travel. A raw numerical value of the slope clearly can not be equated with the cost of overcoming that slope since many different quantifying schemes exist. It may seem reasonable to say that the cost of overcoming a zero degree slope is zero, but it is not necessarily reasonable to say that overcoming a two degree slope is twice as difficult as overcoming a one degree slope. In some way, the numerical value of slope (dh/dx or degrees) must be transformed to a cost.

One example of a transformation from slope to cost is via the hiking function described by Tobler (1993), and developed from empirical data collected by Imhoff (1950). Figure 1 shows the hiking function and derivative cost functions based on time. The function is not symmetric about a slope of zero because the speed at which a person can typically walk is greater on a slightly downward sloping surface than for level ground or an equivalent uphill slope. Maximum speed given this function occurs at roughly 3 degrees of downward slope at 6 kilometers per hour, or 0.16 hours per kilometer. On perfectly level ground, velocity is about 5 kph, and drops to just above 4 kph on a 3 degree upward slope.

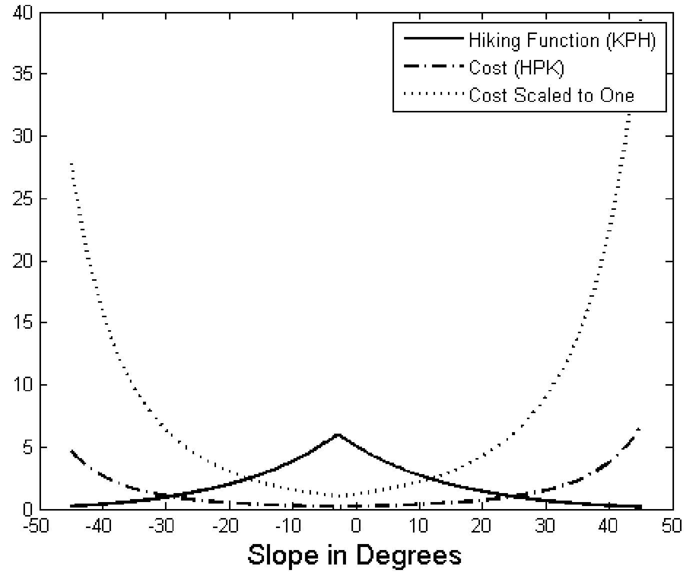

Figure 1 - Tobler's hiking function for adults in hilly terrain and its derivative cost functions. The original function was walking velocity (kph) = 6 { exp -3.5 * abs (dh/dx + 0.05) }. The function is asymmetric about zero slope because it is generally faster to walk downhill than uphill.

Typically, the cost functions derived from speed data are represented as the inverse of the slope/speed relationship. This is correct when solving for time and models utilizing such a cost function will produce least-time routes. Given that time is a significant resource for humans, least-time routes are considered likely candidates when modeling human travel. Whitley and Hicks (2003) used this approach when making GIS-based estimates for likely routes of American Indians through hilly terrain in North Georgia using a terrain sensitive *least-cost path* analysis method pioneered by Warntz (1957). Tripcevich (2007) applied the method to locate probable routes of prehispanic llama caravans in the Andes of South America.

The basic method of least cost path analysis involves multiplying a base cost (in this case distance over ground) by a another cost of movement - a friction cost. In the context of Geographic Information Systems, a raster layer is generally used to create a network where each pixel is a node, and each node is connected to some number of neighboring nodes. This network can then be solved for the least cost path by a variety of shortest-path algorithms. In order to maintain the integrity of relationships between costs, the base cost is scaled [0 Inf] while the friction cost is scaled [1 Inf].

Figure 1 shows the inverse function (hours per kilometer), and the same cost function scaled [1 Inf]. Thus, travel upslope at 30 degrees occurs at approximately 0.67 kph, and takes about nine times longer than at the maximum walking speed. Travel downslope at 30 degrees has a modeled velocity of 0.95 kph and takes about 6.3 times longer than when one travels at the maximum speed.

Crucial to the calculation of slope-based cost surfaces are the initial slope calculations themselves. These are commonly developed from regular raster elevation grids, such as the Shuttle Radar Topography Mission elevation data sets or any other Digital Elevation Model. Zhang, Drake, Wainwright and Mulligan (1999) summarize many of the methods that operate on elevation rasters to return slope rasters. This suite of algorithms typically applies a best-fit surface to each pixel and its eight immediate neighbors. The slope and aspect are then calculated for the surface where it intersects with the central pixel. Other methods of generating slope rasters include assigning each pixel the maximum gradient between the starting pixel and its neighbors, and Horn's (1981) method of calculating third-order finite differences. In each case the output raster of is equal size and resolution to the input raster (excluding edge effects for some methods of calculation), and slopes are always positive.

In contrast, one can calculate slope (dh/dx or atan[dh/dx]) for each *connection* between pixels (Yu, Lee & Nunro-Stasiuk 2003). They may consider directionality, so that slopes may be positive or negative. They may also consider not only immediate neighbors, but also distant neighbors. Links to neighbors immediately up, down, right, and left are known as *rook* connections. Diagonals add *queen* connections, and connections to beyond this are classified as *knight* connections if they are connected in the manner according to their namesake in chess. Figure 2 illustrates different methods of connections for a four square pixel problem. Typical slope algorithms would produce only sixteen slope estimates (one per pixel). Such reduced estimates may be called *point* or *areal* slope measurements, while network-based solutions produce *linear* slope estimates. Previous work analyzing human movement over slopes have been based on areal slope rather than linear slope (Kantner 1996, Whitley & Hicks 2003; Tripcevich 2007). This would have the effect of modeling the agent as one who avoids any slope at all, regardless of whether the actual movement is horizontal or not. Movement along structures like switchbacks, which decrease slope in favor of extra distance, would incur a greater penalty (as a result of the extra distance) than simply minimizing distance on the same slope.

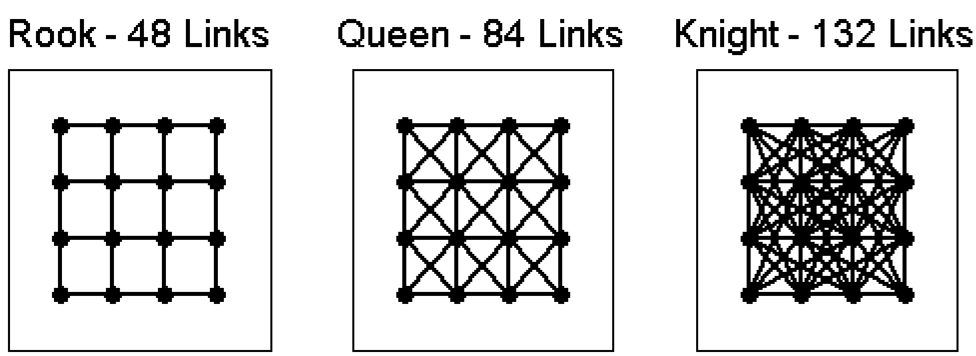

Figure 2 - Network-based slope calculations. Number of links are given for single-direction connections.

There are three reasons why Tobler's hiking function should not be used directly to model costs *as actually applied by humans*. First, downhill speeds, while relatively faster at shallow slopes than their uphill equivalents, are relatively slower at higher slopes. Moving downhill on steep slopes is more dangerous than uphill travel, requires more caution, and typically results in slower speeds. Second, humans have been shown to make systematic errors in estimates of distance, slope, and travel time (Briggs 1973a, 1973b; MacEachren 1990; Proffitt, Bhalla, Gossweiler & Midgett 1995; Freundschuh 1998). These combine to produce patterns of estimates that lead to sub-optimal route choice with respect to time. As a normative model, time-based cost functions derived from empirical relationships between slope and speed will produce time-optimal paths. However, as a descriptive model for how humans actually find routes through hilly terrain due to the impact of slope, inverse hiking functions fall short. Third, slopes impart costs of effort as well as costs of time. A more complete model of human route choice based on slope would include both an expanded set of derivative costs and accommodate cognitive biases in the assessment of these costs.

There are certainly many more drivers to route selection than slope, and the models suggested here are intended to be applied in concert with other drivers, both physical and nonphysical. Slope also acts differently on agents in accordance with their capabilities. Species, age, health, fitness, body characteristics, and experience are only a few additional considerations when considering the capabilities of an agent. For these reasons, the methodologies presented here are designed to be flexible and customizable based on individual agents or kinds of agents. Yet the influence of slope on route choice is so significant that there are bound to be commonalities in the behavioral record of many kinds of agents.

It is important to note the difference between *routes* and *paths* in this context. Routes are a spatial description of how an agent moved in a planned fashion through its environment in a particular instance. Paths are hardened routes or parts of routes, often physically marked in the environment by worn earth or pavement (Montello 2005). Paths are frequently compromises between asymmetric routes. By imposing a common structure, paths reduce resource costs of road construction and cognitive costs of navigation even though the preferred route from A to B and from B to A may well be different. As such, *route finding* models one-way travel and considers the asymmetry of uphill and downhill travel, while *path finding* considers the net cost of uphill and downhill travel on the same segment.

Finally, when the terrain is flat enough, routes and paths are often structured in a way that completely disregards slope. This explains why road building in *relatively* flat land disregards slope in favor of straightness and regularity (as it does in much of the Midwest United States, for example). Effectively, the variation in slope is not worth the additional distance or monetary expense of diverting roads.

2. Alternative explanations of slope costs

Humans tend to overestimate geographic slopes by a surprisingly high margin. Proffitt et al. (1995) found that people overestimate (both verbally and by visual matching) 2° slopes as 10° and 10° slopes as almost 30°. Steep (~30°) slopes had a greater degree of exaggeration from the top (as if moving downhill) than from the bottom. Further, elderly and fatigued subjects tended to overestimate slope to an even greater degree than non-fatigued subjects. The authors conclude that slope overestimation is strongly related to the effort required to overcome it.

Proffitt et al. (1995) fit an exponential model to uphill and downhill slope estimation, the result of which were models indicating the higher overestimation for downhill slopes at high (~30°) values. However, in later work Yang, Dixon and Proffitt (1999) found that individuals tend to greatly overestimate the height of large outdoor features. Given Ross's (1974) observation of the near indistinguishability of extremely high slopes (75° ~ 90°) and the assumption that a 0° slope would be correctly identified, one solution to the slope overestimation problem is that people tend to exaggerate the vertical component to a greater degree than the horizontal component of a slope.

To this end, a simple vertical exaggeration function was fitted to the slope data computed from Proffitt et al. (1995). Figure 2 below shows the original data extracted from Proffitt et al. (*ibid.*), exponential fits for each side separately, and a least-squares vertical overestimation fit (R^2=.98). This analysis indicates downhill slopes are overestimated at approximately 2.3 times the vertical, while uphill slopes are overestimated at 2 times the vertical. This kind of fit has an added benefit over simple exponential fits in that it achieves equality between cognitive slope and true slope at the anchor points of zero and ninety degrees.

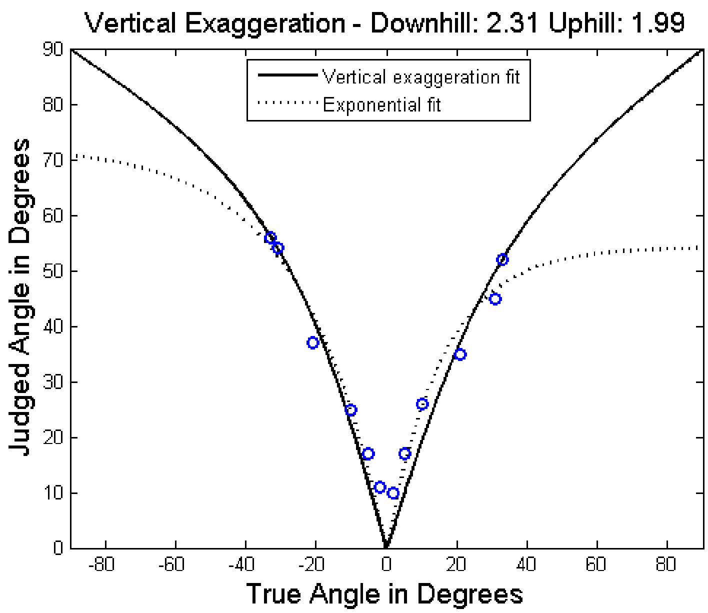

Figure 3 - Model of overestimation of geographic slope, following data from Proffitt et al. (1995). Vertical exaggeration modeling equation: Cognitive slope(deg) = atand{V.E. * (dh/dx) }. Exponential fit: slope(deg) = a{1-exp[b*slope(dh/dx)] }.

If the shape of the curve is indicative of the expected cost of traversal of slope, it represents a stark contrast to the time-based cost curve derived from speed alone. While the time-based cost curve increases slowly over shallow slopes and quickly over steep slopes, costs derived from overestimates of slope do the opposite. Proffitt et al. (1995) argue that this is due to the fact that humans typically navigate over relatively shallow slopes, and thus a higher degree of discernment must occur over these slopes than those which are never traversed.

The slopes on which humans typically operate are astonishingly low for those not accustomed to reviewing the numerical values. For example, California Department of Transportation guidelines specify that no road or driveway exceed 16%, or roughly 9° without special dispensation. Even relatively steep mountain passes rarely exceed 9%, or about 5°. Railway traffic is even more constrained: a review of the charters of American railroads indicates that anything above 3.3% (1.9°) grade is extremely rare and traversal of these requires a special "pusher" locomotive to provide assistance.

Cost surfaces involving slope can be generated in other ways as well. Yu et al. (2003), in a similar least-cost path solution to highway engineering problems, used isotropic average slopes as part of a cost structure for road construction. The authors report that typical cost structures for slope range from 0 (for areas of 0° ~ 3° slope) to 80 (for areas of 12° ~ 16° slope) with slopes over 16° classified as impassable. Such cost structures, however, may not find any viable path between two points in mountainous areas. In practice, switchbacks, bridges, and tunnels are used to accommodate road building in these areas.

3. New models of slope cost

In order to make a comparison to pedestrian travel times described by Tobler, I spent some time driving in the foothills and mountains north of Santa Barbara, California with a GPS unit to find out how my own speeds in a vehicle were affected by slope. Figure 4 shows the distribution of speed against slope over a 2.5 hour, 100 kilometer trip in a vehicle with four wheel drive. Position was recorded with a GPS unit (MTK chipset with WAAS enabled) once per second. Latitude and longitude were converted in Matlab to UTM coordinates via the Mapping Toolbox. These distances were then combined with elevation information directly recorded by the GPS unit to produce slope and speed estimates for each second of the record. These data points are shown on the figure in gray.

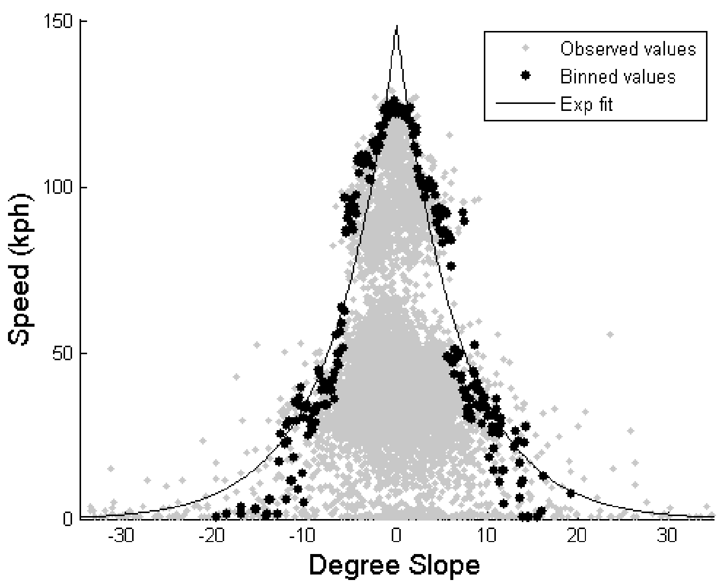

Figure 4 - Distribution of speed against slope over a 2.5 hour, 100 kilometer trip. The exponential fit was nearly symmetrical about zero slope. The final modeled function was velocity (kph) = 150 { exp -8.3 * abs (dh/dx) }.

The record includes a seven kilometer track of freeway travel as well as high quality surface streets in the local foothills and rough unimproved roads in the mountains. As a result, the distribution shows the range of possible speeds for travel at given slopes. Several outliers are present, presumably from GPS error. Additionally, the bounds of the graph are clipped to 35 degrees of positive and negative slope. In fact, only extremely customized vehicles are capable of traveling beyond such slopes, and only then at highly reduced speeds. Slopes greater than 45% (24°) off-road or 60% (31°) on-road are classified by military terrain analysts as impassable by most military vehicles, including tanks (Department of the Army 1990). Although some travel during this record took place on jeep trails in the mountains, it is unlikely that any travel beyond 20° actually took place. These points are likely due to elevation shifts from the GPS record and were excluded from further analysis.

In order to determine a maximum speed function, the maximum speed for each fifth of a degree was computed in the range of -35 degrees to positive 35 degrees. The third highest value in the bin was recorded as the maximum speed for that bin to eliminate outliers caused by GPS error. Figure 3 shows these binned values in black, with an exponential non linear least squares fit in black. The value of the asymmetry parameter was nearly zero (8x10^-4) and the estimated maximum speed parameter approached 150 km/hr (true value 148.7). The resulting scaling parameter produced by nonlinear least squares fitting using these generalized values -8.3.

Travel speeds tend to plateau at the maximum such that for a range of values speed remains equivalent. In automobiles this is partially related to the legal speed limit. However, such plateaus exist for pack animals (Tripcevich 2008) and most likely humans as well. Walking itself is a lower speed of travel than the maximum available to the pedestrian, and represents a compromise between speed, effort expended, and distance expected to travel. Just as an automobile driver may expend more energy to maintain constant speed for 1 degree of positive slope, so might a pedestrian maintain his or her typical walking speed at slight upward slopes by expending more effort.

While a single asymmetric absolute value exponential function may provide a parsimonious description of the relationship between speed and slope, it tends to underestimate speed at shallow slopes and overestimate them at steep slopes. It also is likely to overestimate speeds at steep downhill slopes and underestimate speeds at steep uphill slopes. A splined model that accounts for these bands of slope differently is likely to be more effective, and modern computer modeling techniques make such nuances relatively easy to include in a GIS-based model.

A second method of assessing the cost of slope involves the construction of a probability density function that described the slope values along an existing trail. To this end, a 12.5 km stretch of mountain trail and road used by hikers and experienced equestrians was selected in the mountains north of Santa Barbara, CA to serve as an example. The eastern portion of the path is Arroyo Burro Road, a rough semipaved mountain road that generally follows a ridge. The western portion is the upper length of Arroyo Burro Trail, a somewhat more difficult mountain trail that generally follows a valley. Both paths are highly influenced by terrain, contain a number of switchbacks, and do not cross any water bodies that are a barrier to movement.

The entire path segment was digitized by inspection from a 1 meter resolution color DOQQ available from the USGS Seamless Data Server and with the reference of several published hiker GPS waypoints and guides. Elevation data was acquired from the National Elevation Dataset at a resolution of 1/3 arc second (~10 meters) for the surrounding area (-119.60/-119.52;34.68/34.63). Figure 5 below shows the paths and a hillshaded image generated from the elevation data set to highlight the local terrain.

The digitized trails consisted of a total of 852 segments with a mean length of 14.7 meters. Given the relatively high resolution of the path data, linear slope calculations used a bicubic interpolation of the underlying elevation data to produce an estimate of elevation for each point on the path.

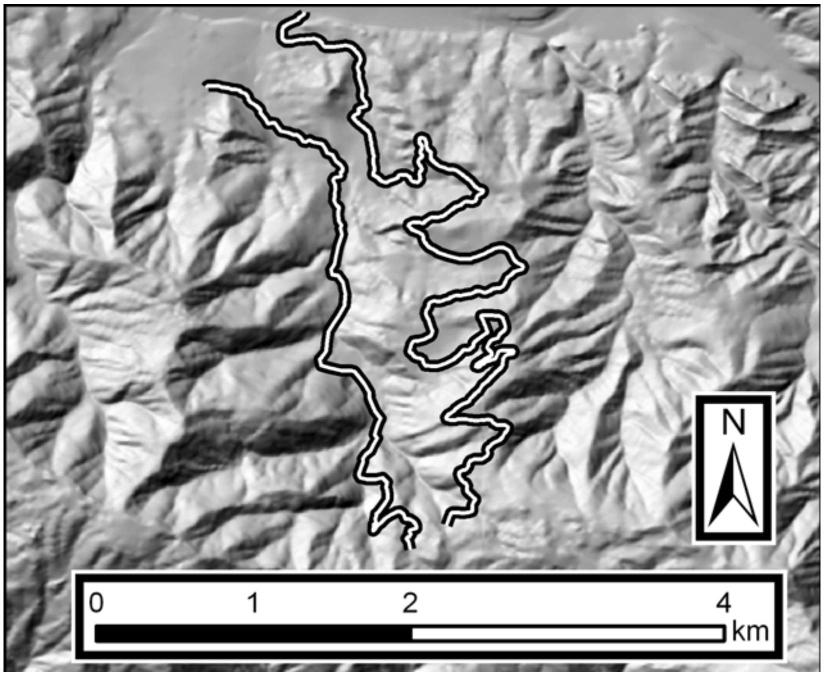

Figure 5 - Horse and human recreational hiking trail in the mountains north of Santa Barbara. The eastward path is Arroyo Burro Road, an unimproved jeep trail. The westward path is a segment of Arroyo Burro Trail, used by horses and hikers.

Since distances between points varied, each segment's slope was weighted according to its associated distance, and each value and its additive inverse were included in the final distribution since the path could be traversed in either direction. Figure 6 shows the resulting estimated functions for the probability density function (pdf). A maximum likelihood estimation fit of the normal distribution for the *symmetric* data yielded estimates of a mean of zero and variance of approximately 11.9 degrees. An exponential fit resulted in an estimated scaling factor of 9.2. In this case, neither fit was particularly close to the kernel density estimate, though the general behavior was modeled well.

The translation of an observed probability distribution to a cost function is not straightforward, but there are good reasons to think that it may obtain promising results in terms of route prediction. It is worthwhile to note that the assumption of the translation - that things observed in greater frequency are less costly than those observed in lesser frequency - only holds as long as there is equivalence of some other sort between classes. In common consumable goods this typically means an equivalence in price, quantity available, or both. In the case of slope as a driver or route selection, any point may be traversed at any slope between the maximum point (or areal) slope and its additive inverse. Since we know *a priori* that human prefer shallower slopes to steeper ones, the requirement of equivalence is met.

The method of transforming a probability density function to a cost function is easiest to understand through an example. Consider the following case, in which I am presented at a luncheon with a basket containing several different types of fruit. My selection of fruit is likely to show patterns of consumption according to my tastes. If I attend such a luncheon every day, over time we may observe that I consume 80 apples for every 19 oranges for every 1 peach. In some sense apples are worth nearly four times as much as oranges, and eighty times as much as peaches. This is nothing more than a common utility function described by von Neumann and Morgenstern (1944) during the early development of game theory.

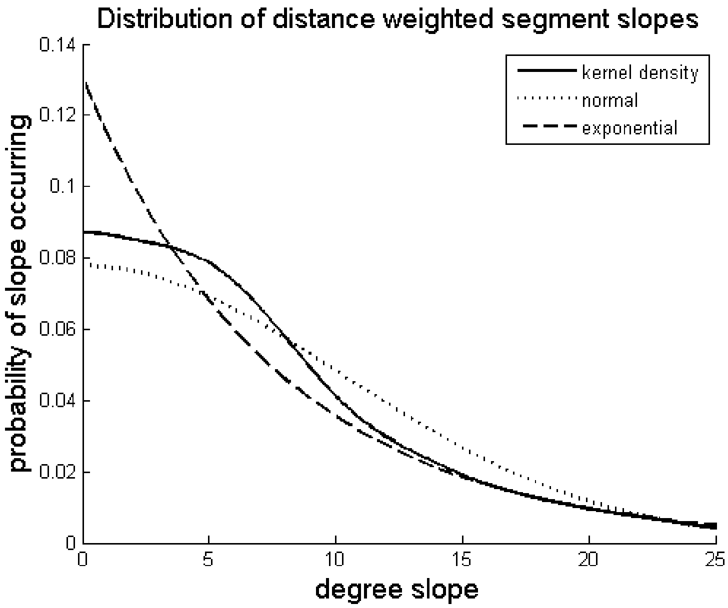

Figure 6 - Probability density distribution estimates for Arroyo Burro Trail and Road (combined).

We can also frame these choices in terms of costs. To eat a peach is then eighty times as costly as the base unit of consumption - an apple - and an orange is nearly four times as costly. From a univariate standpoint, the peak of the pdf represents the most frequently chosen and least costly alternative. All other points are scaled to this such that the resulting distribution is given in terms of its highest probability - in other words, scaled [1 Inf]. The costs are then multiplied by the distance over which the cost is applied, resulting in the overall cost surface.

Figure 7 below show three hypothesized cost relationships for paths traversing slopes zero to twenty degrees. Since we are considering the traversal of paths (which may be taken in either direction), we must adjust the probabilities derived from the distribution fits accordingly. In the case of Tobler's hiking function, the mean rate (hours-per-kilometer) was taken for the range of slopes above for both positive and negative values. Relatively minor differences in the pdf fits for the exponential and normal models result in quite large differences over shallow slopes. While the exponential model suggests costs begin to increase quickly, the normal model and hiking function derived costs suggest that costs are not significantly different from zero until the slope reaches about ten degrees. The hiking function, which approximates only the time cost involved in overcoming slope, remains relatively flat before rising sharply at about twenty degrees. The cost function derived from the kernel density estimate of the probability density function features the same low costs for shallow slopes as the normal pdf and hiking functions, but costs accumulate more quickly than either of these at higher slopes. Table 1 lists the parameter estimates for the exponential and normal cost structure for the trail and road separately.. Note that the trail cost function is somewhat shallower (i.e., more permissive with respect to slope) than is that of the road.

The final step in assessing which cost structures most closely model human behavior is to use the function as a selection algorithm in the same space. The more closely the route selected by the algorithm mirrors the routes actually chosen by humans, the more likely it is that our hypothesized cost structure is correct.

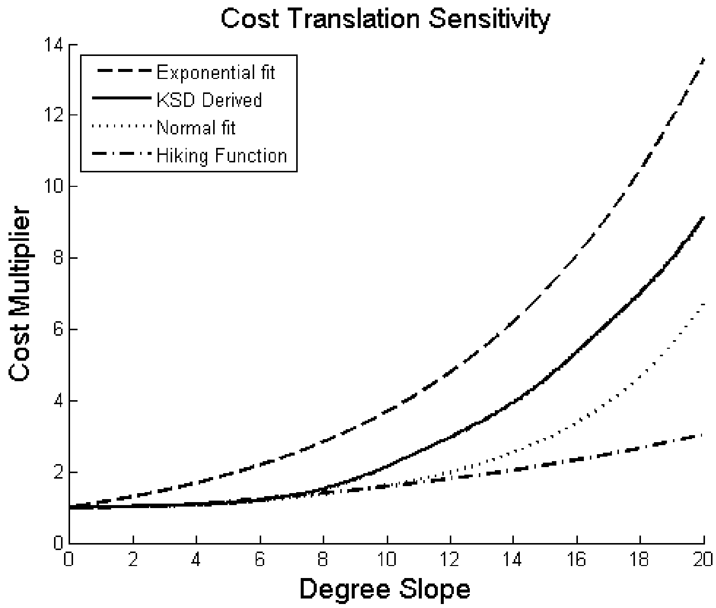

Figure 7 - Translated cost functions for probability density functions from Figure 4 and absolute value hiking function for mountain trail and road combined.

Scale Parameter	Trail	Road
Norm	11.92	9.55
Exp	9.26	7.01

Table 1: Parameter estimates for trails and roads.

4. Model evaluation

In order to evaluate the cost structure, the previously mentioned 1/3 arc second elevation raster was transformed to network or graph structure. The center of each pixel became a node, and each node was connected to its eight neighboring nodes (*i.e.*, the queen's pattern). Initial tests showed very little difference at these high resolutions between a network structure consisting of queen's connections (8 connections) and knight's connections (16 connections). The structure was developed in such a way that each link could be traversed in only a single direction. In this way, the structure allows for the development of route models that explicitly take into account the different costs involved in traveling upslope versus downslope. Although only path results are presented here, we successfully tested the application of asymmetric slope transformation functions as they would apply to routes.

Each link was assigned two costs: distance and slope cost. Distance was great circle distance (in meters) calculated via Matlab's Mapping Toolbox for the GRS80 ellipsoid. The slope was calculated by taking the difference of the elevation for the two nodes that defined each link. Each slope was then transformed using the pdf-derived cost mechanism described above.

The exponential transform was used (and the associated values in Table 1) as extensive testing showed that even transforms derived from the kernel density function were too permissive and resulted in too steep of slopes selected for by the algorithm. A Java implementation of Dijkstra's (1959) Algorithm was used to solve the graph for the least cost path.

Figure 8 shows the paths selected for Arroyo Burro Trail and Road, each based on its respective exponential pdf-derived cost function. Both modeled paths follow the general trend of the true path they attempt to emulate. The trail model was somewhat closer in overall path as well as distance. The road model was correct to within 20 meters over the first 750 meters of the path, and to a similar margin of error for the last kilometer. The middle sections do not match well spatially, but visual inspection of the modeled path shows it is as viable with respect to slope traversal as is the true one.

An analysis of the slopes selected along the modeled paths indicate a higher proportion of slopes in the 5 to 10 degree range than actually occur. However, over a longer distance (15 km) the distributions underlying the cost structure and the final distance weighted slopes observed on the modeled path were quite similar. Figure 9 shows these distributions. The selection distribution in this case was that derived from the Arroyo Burro Trail (exponential fit with scaling parameter of 9.26). Over the longer path, slopes were actually biased in favor of very low slopes (0 to 4 degrees).

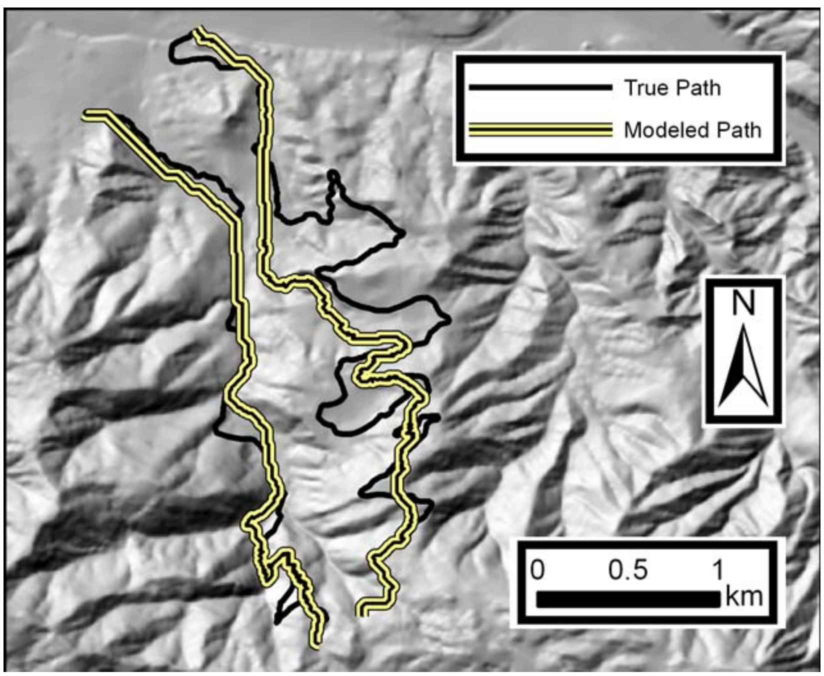

Figure 8 - Modeled paths for Arroyo Burro Trail (left) and Road (right).

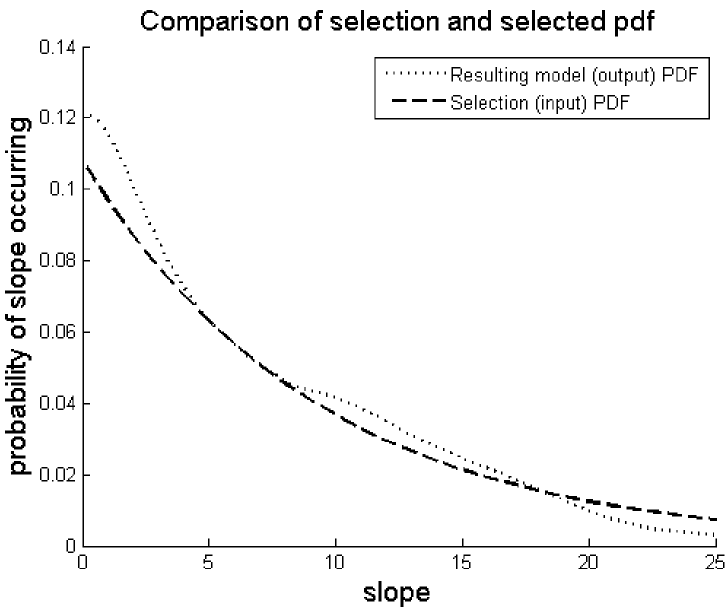

Figure 9 - The pdf underlying the cost structure compared to the pdf of distance weighted slopes along selected path.

For the purposes of comparison, two endpoints (11.2 km apart) were selected. The model was for two different cost structures: the pdf derived exponential cost function, and the hiking function. Both selected visually reasonable paths. As expected, the hiking function path was shorter (15.3 km) and the mean slope was steeper (6.93 degrees). The pdf derived function was much longer (21.4 km) but traversed shallower slopes on average (6.01 degrees). Figure 10 shows each of these paths.

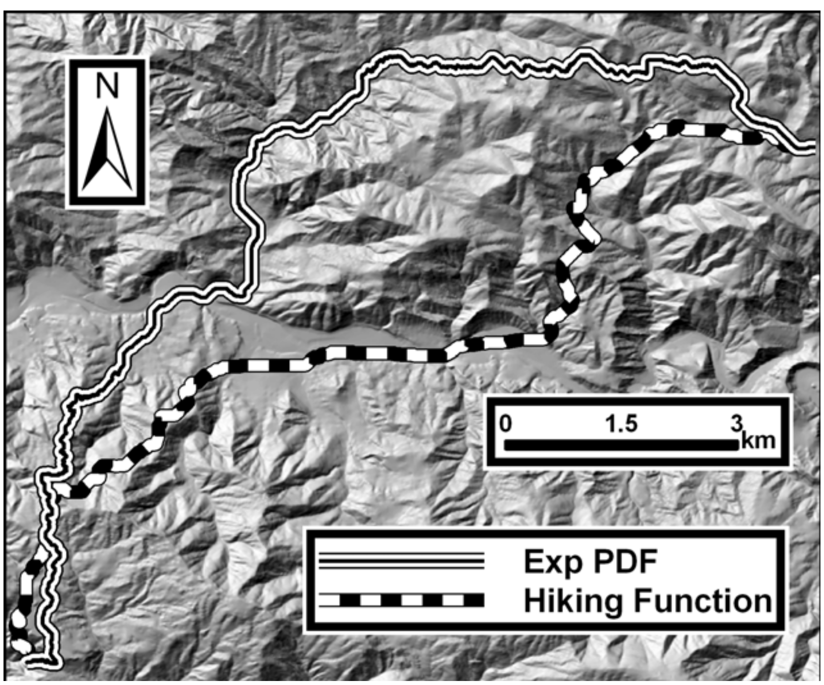

Figure 10 - Underlying cost structures produce very different routes.

In order to test how the linear slope model would perform over a much larger distance, we attempted to find a viable route between Hirat and Kabul in Afghanistan, which are a distance of approximately 640 kilometers apart. British foreign service expert Rory Stewart spent nearly a month traveling between these two cities on foot, and detailing the account in *The Places In Between*. Several maps published in the book along with descriptions in the text were used in combination with road and settlement data provided by the Afghanistan Information Management Services to map the locations of the settlements through which Stewart passed.

Stewart describes two viable routes from Hirat to Kabul. The first is long detour around the high central mountains, through Khadahar and along Afghanistan's highway A01, while the second travels through the central mountains, largely along highway A77. It is much shorter, but more difficult in terms of slope traversal. Many of the vehicle roads in the high passes of the central mountains are closed during the winter months, making the various villages accessible only by air or on foot. Both routes have been well used historically by tradesmen and invaders.

Figure 11 - Model of Rory Stewart's walk described in The Places In Between. Stewart walked from Hirat to Kabul along the northern route in the figure stopping at waypoints indicated.

Figure 11 shows the results of the model along with waypoints through which Stewart passed during his trip. The initial model used an inverse exponential cost model with a scaling factor of 7. This model followed the true trail quite well until Barakhana, where it began to follow a southerly trail that bypassed Bamyan. A scaling factor of 3 ultimately produced the closest results, though it still diverted south through Dowlat Yar. Interestingly, it is at this point that the highway (A77) also diverts south through Moshi on its way to Kabul. The northern route, through Bamyan, is largely devoted to foot and animal travel. When Bamyan was included as an intermediary point in the model, Stewart's path was picked up extraordinarily well. This path rejoined vehicle highway (A77) west of Kabul.

At several points in the account, Stewart describes the southern route as being flatter and easier, though longer. With a scaling factor of 2 on the inverse exponential cost function, the model correctly identifies the southern route through Kandahar.

This large-scale model used reduced resolution (~1 km) elevation data due to the high demands on memory that would otherwise be incurred by modeling such a great extent. In this and other tests of the model we found that over very large distances this reduction in resolution did not seem to affect performance - in fact in most cases at this scale, the use of more aggregated data tended to *increase* performance. This should not be too surprising from a cognitive perspective. Humans tend to aggregate features in their environment into useful chunks. It is likely that in development of these routes humans tended to think more broadly and worry about the finer scale issues of slope traversal only when immediately confronted with it.

5. Discussion

Although humans choose routes based on more than slope, changes in elevation tend to exert a powerful influence on route selection in hilly and mountainous terrain. The model presented here was quite adept at picking out existing roads and reasonable routes at varying scales and resolutions. This model combines the cost associated with linear slope traversal described by Tober (1993) and Yu, Lee and Nunro-Stasiuk (2003) with an enhanced cost function supported by cognitive research (Yang, Dixon, and Proffitt 1999) and developed from the distribution of slopes derived from paths used by humans. The methodology allows for the easy construction of slope-cost functions, since the inverse exponential scaling factor is exactly equal to the mean slope that an agent travels, and can thus be easily revised based on the capabilities of the agent. Additionally, the method can be adapted to use entirely empirical (kernel density descriptions) of slope traversal behavior. As with all aggregate surface friction models, slope costs can be combined with other elements to include other movement costs as well (e.g., land cover, water, etc.).

While linear slope is dominant among costs associated with changing elevation, other elements of slope cost can be modeled as well. Most GIS packages allow the computation of point or areal slope, and these estimates of maximum elevation change at a single location matter for how humans or other animals move. Even a linear walk at zero degrees on a side slope of forty-five degrees would be extremely difficult. Curvature matters as well: while ridge and valley roads are both common, ridges tend to be relatively more common. As these factors are included, it is likely that even better results can be obtained.

In the models presented here, we tended to regard uphill and downhill slopes as equally costly with respect to paths, since paths can be traversed in either direction. However, when unidirectional route planning is considered explicitly, asymmetries in human travel behavior identified by Tobler (1993) and verified by psychological research (Yang, Dixon and Proffitt 1999) should be included. The pdf-derived cost method is easily adapted for this as well, simply by extending the function to negative as well as positive slopes.

6. Conclusions

Existing modeling techniques that utilize slope as a driver for least cost path solutions involving human and animal improvement would be much improved from a predictive point of view if they reproduced the relevant biases that exist in the agents they attempt to model. We have shown that a much steeper cost structure than time alone improves the path finding ability of friction-surface based approaches to mobility modeling.

Such techniques are of clear use to archaeologists who might use such an analysis to hypothesize travel pathways and focus archaeological surveys along highly probable routes. Given a large enough set of modeled pathways, trade settlements might be identified from crossroad locations. Even in areas were there are no known settlements, multiple endpoints could be chosen within an area to find likely travel corridors.

Heth and Cornell (1998) have successfully used GIS to aid in the search for persons lost in wildness areas largely by focusing the search to likely areas. Mobility analysis using slope is very likely to further restrict the likely areas of travel by focusing attention on travel corridors.

Given the global availability of high-resolution spatial elevation data, one of the next steps of this project is to automate the analysis process and to develop classes of movement so that empirical analysis of existing pathways is less necessary. Side slope and curvature values must ultimately be included in every friction surface involving slope. Edge finding image processing algorithms on satellite imagery may also help identify tracks and trails that are not well mapped, but which may remain behind. Slope is such an important driver to route selection that effective mobility modeling based on empirical data should be a high priority, especially given the wide variety of applications.

Acknowledgements

The author is most grateful to Keith Clarke, Waldo Tobler, Harry Starr, and Nicholas Tripcevich for their valuable comments and suggestions on early drafts of this paper.

References Cited

Bailenson, J. N., Shum, M. S., & Uttal, D. H. (1998). Road climbing: Principles governing asymmetric route choices on maps. J*ournal of Environmental Psychology, 18,* 251-264.

Bailenson, J. N., Shum, M. S., & Uttal, D. H. (2000). The initial segment strategy: A heuristic for route selection. M*emory & Cognition, 28*(2), 306-318.

Briggs. (1973a). Urban Cognitive Distance. In R. Downs & D. Stea (Eds.), *Image and Environment: Cognitive Mapping and Social Behavior* (pp. 361-388). Chicago: Aldine.

Briggs. (1973b). On the relationship between cognitive and objective distance. In W. F. E. Preiser (Ed.), *Environmental Design Research* (Vol. 2, pp. 182-192). Stroudsburg, PA: Dowden, Hutchinson and Ross.

Dijkstra, E. W. (1959). A note on two problems in connection with graphs. N*umerische Mathmatik, 1*, 269-271.

Freundschuh, S. M. (1998). The Relationship Between Geographic Scale, Distance and Time as Expressed in Natural Discourse. In M. J. Egenhofer & R. G. Golledge (Eds.), *Spatial and Temporal Reasoning in Geographic Information Systems* (pp. 131-142). New York: Oxford University Press.

Golledge, R. G. (1995, May). *Defining the Criteria Used in Path Selection.* Paper presented at the International Conference on Activity Scheduling, Eindhoven, The Netherlands.

Heth, C. D., & Cornell, E. H. (1998). Characteristics of Travel by Persons Lost in Albertan Wilderness Areas. *Journal of Environmental Psychology, 18*(3), 223-235.

Hochmair, H. (2004). Towards a classification of route selection criteria for route planning tools. In P. Fisher (Ed.), *Developments in Spatial Data Handling* (pp. 481-492). Berlin: Springer.

Kantner, J. (1996, April). *An Evaluation of Chaco Anasazi Roadways*. Paper presented at the 61st Society for American Archaeology Annual Meeting, New Orleans, Lousiana.

MacEachren, A. M. (1980). Travel time as the basis for cognitive distance. *The Professional Geographer, 32*(1), 30-36.

Montello, D. R. (2005). Navigation. In A. Miyake & P. Shah (Eds.), *Cambridge handbook of visuospatial thinking* (pp. 257-294). Cambridge: Cambridge University Press.

Proffitt, D. R., Bhalla, M., Gossweiler, R., & Midgett, J. (1995). Perceiving geographical slant. *Psychonomic Bulletin & Review, 2*(4), 409-428.

Ross, H. E. (1974). *Behaviour and perception in strange environments*. London: George Allen & Unwin.

Stahl, C. W. (2005). *Accumulated Surfaces and Least-Cost Paths: GIS Modeling for Autonomous Ground Vehicle (AGV) Navigation.* Virginia Polytechnic Institute and State University, Blacksburg, VA.

Stewart, R. (2006). *The places in between* (1st U.S. Ed.). Orlando: Harcourt, Inc.

Tobler, W. (1993). N*on-isotropic geographic modeling* (Technical Report No. 93-1). Santa Barbara, CA: National Center for Geographic Information and Analysis.

Tripcevich, N. (2007). Q*uarries, Caravans, and Routes to Complexity: Prehispanic Obsidian in the South-Central Andes. U*npublished Dissertation, University of California Santa Barbara, Santa Barbara, CA.

Tripcevich, N. (2008, March). *Llama caravan transport: A study of mobility with a contemporary Andean salt caravan.* Paper presented at the 73rd Annual Meeting of the Society for American Archaeology, Vancouver, B.C., Canada.

Department of the Army. (1990). *Terrain analysis*. Washington, DC: Department of the Army.

Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior.* Princeton, NJ: Princeton University Press.

Warntz, W. (1957). Transportation, Social Physics, and the Laws of Refraction. *Professional Geographer, 9*(4), 2-7.

Whitley, T. G., & Hicks, L. M. (2003). A Geographic Information Systems Approach to Understanding Potential Prehistoric and Historic Travel Corridors. *Southeastern Archeaology, 22*(1), 77-91.

Yang, T. L., Dixon, M. W., & Proffitt, D. R. (1999). Seeing big things: Overestimation of heights is greater for real objects than for objects in pictures. *Perception, 1999*(28), 445-467.

Yu, C., Lee, J., & Nunro-Stasiuk, M. J. (2003). Extensions to least-cost path algorithms for roadway planning. *International Journal of Geographic Information Science*, 17(4), 361-376.

Zhang, X., Drake, N. A., Wainwright, J., & Mulligan, M. (1999). Comparison of slope estimates from low resolution DEMs: Scaling issues and a fractal method for their solution. *Earth Surface Processes and Landforms*, *24*, 763-779.