

# Modelling potential environmental impacts of science activity in Antarctica

David O’Sullivan

Fraser J. Morgan

## Abstract

We use GPS data collected on a science expedition in Antarctica to estimate hiking functions for the speed at which humans traverse terrain differentiated by slope and by ground cover (moraines and rock). We use the estimated hiking functions to build weighted directed graphs as a representation of specific environments in Antarctica. From these we estimate using a variety of graph metrics—particularly betweenness centrality—the relative potential for human environmental impacts arising from scientific activities in those environments. We also suggest a simple approach to planning science expeditions that might allow for reduced impacts in these environments.

## 1 Introduction

Overview of Antarctic science: when and where, its intensity etc. Background on international treaties, etc.

Relevant findings as to human impacts in Antarctica. Note that in this environment even ‘leave only footprints’ is likely impacting the environment in significant ways.

Overview of sections ahead.

## 2 Our approach and related work

We chose to explore the question of where human impacts are likely to be strongest using an approach closely related to work on patterns of human movement in archaeology (Verhagen et al. 2019) where likely and potential movement paths of humans across landscapes have been used to infer the settlement structure and human geography of large-scale landscapes. Closely related work in biology investigates the structure and geography of animal transportation networks (Perna and Latty 2014). Both approaches rely on the idea that humans or animals move around in an environment in time or energy efficient ways. These approaches rely on hiking functions that relate speed of movement across a terrain to its slope.

Hiking functions must be applied in some context where locations across a landscape are connected to one another. Because hiking functions are asymmetric, with estimated speed of movement up slopes different than estimated speeds down the same slope, landscape must be represented in a way that allows for this asymmetry. We therefore represent terrain in our landscapes as directed graphs (or network) of connections between locations regularly distributed in planimetric space across the landscape of interest. Because the graphs are directed the costs associated with movement between two locations can be different depending on the direction of movement. Additionally, we associate with locations (i.e., vertices in the graph) the ground cover at the location, which also affects the speed at which it can be traversed. Because the ground cover in the Antarctic environments under study can be broadly categorised into only two types, moraine and rock, we use the ground cover of a location to switch between two estimated hiking functions, rather than the more widely used approach of penalising movement on different ground covers by applying cost factors. We consider previous work on hiking functions and directed graphs in more detail below.

### 2.1 Hiking functions

Prisner and Sui (2023) provide an overview of a variety of functions that have been used to model how hiking times and speeds vary with terrain slope. They consider longstanding rules of thumb (Naismith 1892), and

later modifications thereof (Langmuir 1984), along with more recent such guidance from the Swiss and German Alpine Clubs (Winkler et al. 2010; Deutscher und Österreichischer Alpenverein 2011). These functions estimate the time taken to travel 1km, referred to as *pace*, based on slope expressed as *rise over run*, that is change in elevation divided by horizontal distance. They are all piecewise functions with sharp changes in estimated pace at specific slopes.

Alongside these hiking pace functions Prisner and Sui (2023) also present hiking speed functions (generally referred to as simply *hiking functions*) from Tobler (1993 generally considered the first hiking function) and more recent, related but more firmly empirically grounded alternatives offered by Márquez-Pérez et al. (2017), Irmischer and Clarke (2018), and Campbell et al. (2019). Another hiking function not discussed by Prisner and Sui (2023) is presented by Rees (2004). These hiking functions are all continuous in the slope of the terrain so that  $v = f(\theta)$ , where  $v$  is the speed, and  $\theta$  is the slope. They can all be parameterised to control the maximum speed attainable, the slope at which maximum speed is attained (expected to be a shallow downhill slope), and the rate at which speed falls off with increasing slope.

The functional form of some of these functions is shown in Table 1 and graphed in Figure 1.

Table 1: Functional forms of hiking functions

Description	Equation	Examples
Exponential	$ae^{b \theta-c }$	Tobler (1993), Márquez-Pérez et al. (2017)
Gaussian	$ae^{-b(\theta-c)^2}$	Irmischer and Clarke (2018), Campbell et al. (2019)
Lorentz	$\frac{a}{[b+d(\theta-c)^2]}$	Campbell et al. (2019)
Quadratic	$a + b\theta + c\theta^2$	Rees (2004)

The parameterisations of the functions in Figure 1 have been chosen for illustrative purposes only, although the parameter values for the exponential

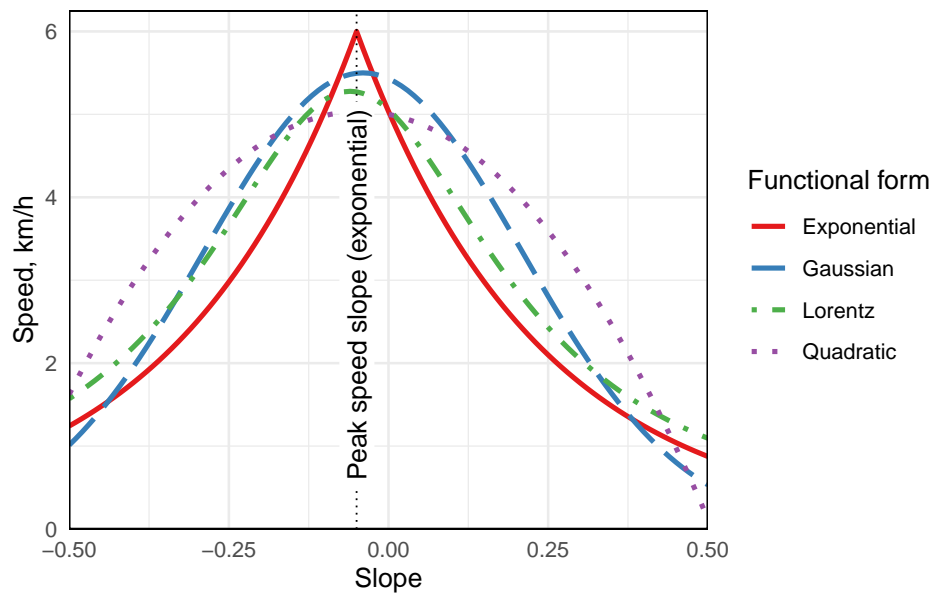


Figure 1: Example hiking functional forms: Exponential (Tobler 1993), Gaussian (Irmischer and Clarke 2018), Lorentz (Campbell et al. 2019), and Quadratic (Rees 2004).

function shown are  $v = 6e^{-3.5|\theta+0.05|}$  as suggested in Tobler (1993). These parameters give *Tobler’s hiking function* but, as has been noted elsewhere, (see Herzog 2010; Campbell et al. 2019), are based on a poorly specified fit to secondary data on hiking speeds presented by Imhof (1950, 217–220). Nevertheless these parameter values are widely applied in the literature.

All these hiking functional forms are somewhat *ad hoc*. They exhibit desirable, expected properties: a peak speed at a slope near zero, which we expect to be slightly negative (i.e., downhill), and continuously falling speeds as the slope deviates from the slope of peak speed. However, there is no theoretical basis for the specific functional forms listed in Table 1. More principled approaches might be developed based on the literature on the physiology of human movement, see e.g. **CITATIONS NEEDED HERE**. In general, approaches based on minimising energy use yield similar results to empirical speed-slope functions, although it is worth noting that they more reliably generate zig-zag or ‘switchback’ movement behaviour on steep slopes (Llobera and Sluckin 2007). However, these are hard to implement, and we have adopted empirically-derived and locally-specific hiking functions following Márquez-Pérez et al. (2017), Irmischer and Clarke (2018), and Campbell et al. (2019). This choice is based on available data and the goals of our study, where the relative cost of different potential routes in the landscape is more important than exact prediction of routes. In practice, almost any function with the peaked form of those shown in Figure 1 is suited to our requirements.

It is commonplace in many applications to also incorporate a penalty on movement contingent on landcover, especially for off-track or off-road movement. For example the speed attainable off-track in forested terrain might be only half that attainable in grasslands. Unsurprisingly, there are no widely agreed landcover penalties, but see for example, those compiled by Herzog (2020). A hiking function derived by Wood et al. (2023) includes the local gradient of the terrain (i.e., the maximum slope at each location, not the slope in the direction of movement) as a covariate in the estimated function. It is possible that this kind of approach treating landcover as a categorical covariate in the estimation of more complex hiking functions might be more widely applied in future work. In our application, because there are only

two kinds of navigable landcover—moraine (or gravel) and rock—we chose instead to estimate two hiking functions, one for each landcover, and estimate movement costs conditional on the landcover present at each location. This also has the advantage of allowing for differing effects of slope on speed due to landcover, where for example, gravel might allow more rapid movement on the level surfaces, but more rapidly reduce speed on slopes. Details of this approach are reported and discussed in Section 4.1.

## 2.2 Representing the landscape

Hiking functions are usually asymmetric with respect to slope, with the highest attainable speed at a slight downhill slope. This asymmetry means that unless analysis is focused on assessing movement cost from a single origin or to a single destination, it is necessary to represent the landscape in a way that can accommodate asymmetry. We therefore represent the landscape as a directed graph. Graph vertices are locations with associated elevation and landcover. Graph edges are directed links between vertices for which a change in elevation between the start and end vertex can be calculated, and a slope derived, based on the elevation difference and horizontal distance between the vertices.

The most obvious way to derive such a graph is to assign each cell in a raster digital elevation model (DEM) to a graph vertex, so that graph vertices are arranged in a regular square grid or lattice. This approach has been widely adopted **CITATIONS NEEDED HERE** However, this still requires a decision to be made regarding the definition of graph edges, which can be thought of as *allowed moves* between vertices. Three possibilities on a square lattice—Rook’s, Queen’s, and Knight’s moves—are shown in Figure 2. Another option shown in the figure is to lay out a regular hexagonal grid of vertex locations and link each vertex with its six nearest neighbours.

The advantage of a square lattice is the ease of deriving vertex elevations assuming a DEM at the desired lattice spacing is available. A hexagonal lattice on the other hand will usually require that elevation values be interpolated from a DEM to the lattice locations.

Both square and hexagonal lattices can lead to geometric artifacts when

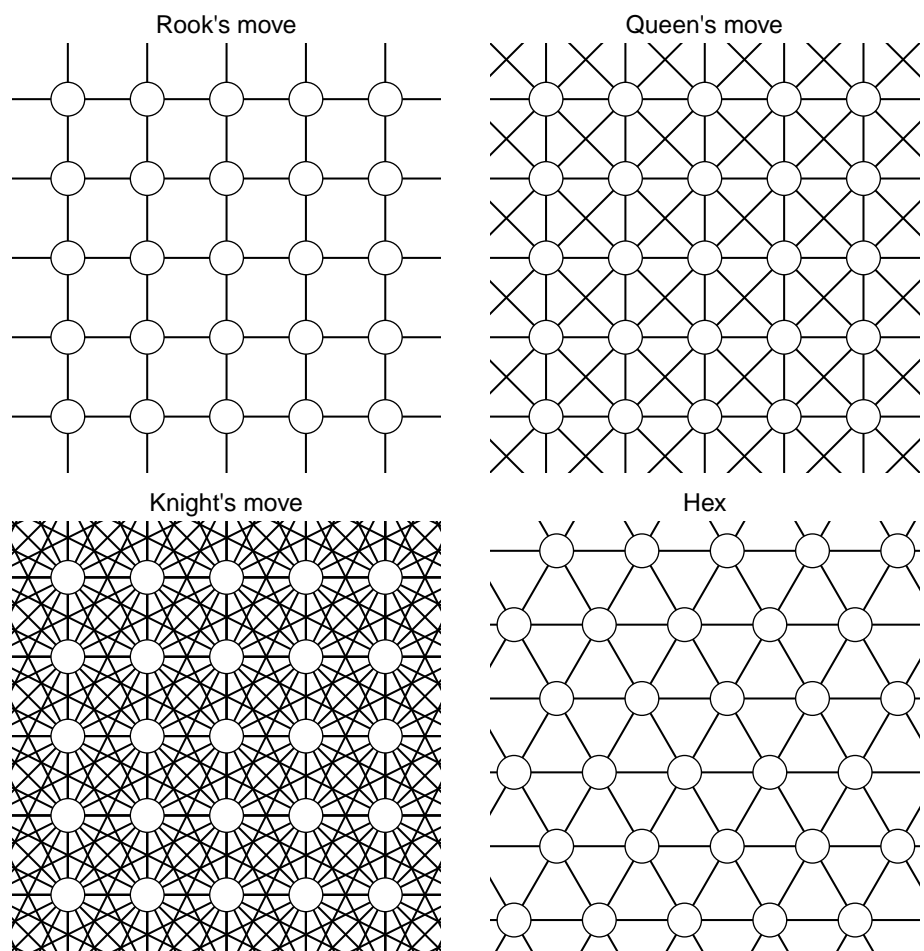


Figure 2: Possible graph lattices.

the allowed moves are used to determine contours of equal travel time or *isochrones* from an origin point. On level terrain the Rook’s move produces diamond-shaped isochrones, while the Queen’s move produces octagons. In general isochrones will be polygons with as many sides as the number of neighbours of each vertex. While the 16-gons of the Knight’s move lattice are likely to be close enough to circular for many purposes, the resulting graph is much larger than the alternatives. Etherington (2012) suggests that by combining results from randomly generated planar graphs such geometric artifacts can be removed, but this substantially increases the computational requirements. Another computationally intensive approach might draw on methods for estimating geodesics on complex triangular meshes (Martínez et al. 2005). Because the geometric effects of the graph structure are masked when we introduce varying movement speeds due to the slope of each edge, we do not consider these more complex approaches necessary in our application. Nevertheless, more circular base isochrone shapes are to be preferred, all other things equal.

A further complication when a lattice includes edges of varying length as in the Queen’s and Knight’s move lattices, is bias in the estimation of slopes leading to lower estimated movement costs for longer edges. This is because longer edges (e.g., the knight’s moves) may ‘jump’ across intervening segments of varying slope, smoothing them to a single slope estimate. Such smoothing will usually result in estimated speeds along that edge higher than those that would accumulate along the intervening sections of varying slope. This problem is discussed in a slightly different context by Campbell et al. (2019, 96–98).

Any choice of graph structure is necessarily a compromise, and estimates of movement rates are always an approximation (Goodchild 2020). We consider the differing edge lengths introduced by all the square lattices other than the Rook’s case to be problematic, and favour the hexagonal lattice structure over the simple square lattice because a base hexagonal isochrone is preferable to a diamond shape.

Having settled on a hexagonal lattice for the graph structure, at a chosen resolution we set out a regular hexagonal grid of locations across the study



area, and assign to each point an elevation by interpolation from a DEM. Because the DEM is at finer resolution than the hexagonal lattice, the choice of interpolation method is not a major concern. Our approach applies a bilinear interpolation method based on elevation values in the DEM cell the graph vertex falls in and its four orthogonal neighbours. Based on the difference in elevation of the vertices at each end of each edge we estimate a slope, and also a traversal times using our estimated hiking function. The important point here is that our graph is *directed* so that different traversal times  $t_{ij}$  and  $t_{ji}$  are estimated for moving between locations  $i$  to location  $j$  depending on the direction of movement.

As noted in the previous section, we estimate two hiking functions, one for moraine and one for rock landcover. When an edge connects locations of two different landcover types the estimated traversal time is the mean of the traversal times for each landcover.

### 3 Data sources

#### 3.1 Antarctic geospatial data

Cox et al. (2023b), Cox et al. (2023a), Felden et al. (2023)

[FIGURE showing location information etc., based on these data]

#### 3.2 GPS data from an expedition

Devices used, and associated protocols for scientists while on site.

[FIGURE showing GPS traces in context]

Initial filtering to remove implausibly high-speed movement fixes – these largely related to helicopter movements and arrival/departure from the valley.

Additional cleanup of GPS data required: filtering to remove many GPS fixes with associated low movement rates associated with time spent at base camps, rest stops, or experimental sites.

Crude filters: short distances between fixes, large turn angles between fixes,

and seemingly anomalous high speeds recorded at 0 slope angles (yet to be explained but required to obtain reasonable results).

Density-based filtering: hexbinning of fixes to identify non-purposive movement areas.

[FIGURE showing filtering steps and impact on slope vs speed scatter-plot/boxplots]

## 4 Methods and results

### 4.1 Hiking functions

Fit three alternative functional forms: exponential Tobler (1993, pp 1–4, based on data from Imhof 1950), Gaussian Irmischer and Clarke (2018), and Student’s  $t$  distribution, the last of these as a heavier-tailed alternative to Gaussian. Other approaches rely on rule-of-thumb hiking time calculations such as...

The Gaussian form offered the best overall fits to our data. Although Tobler’s exponential form is much used it is somewhat implausible that hiking speeds would peak sharply at a very specific slope angle, and so we adopt the Gaussian form.

[FIGURE showing hiking function fits about here]

In previous work researchers have applied a ground cover penalty cost to a base hiking function to estimate traversal times. In our case there are effectively only two terrain types, moraines (gravel) and bare rock. This makes the option of fitting different functions to each attractive and reasonable, and the results of fitting distinct functions to each support this approach. The peak speed on bare rock is attained on steeper downhill slopes than on gravel, perhaps indicative of the greater care required on downhill gravel slopes. Meanwhile the highest speeds on level terrain are attained on gravel. Our fitted functions are shown below.

[FIGURE comparing moraine and rock hiking functions about here]

## 4.2 Landscapes as graphs

We developed R code (R Core Team 2024) to build graphs (i.e. networks) with hexagonal lattice structure and estimated traversal times for graph edges derived from our hiking functions. Graphs are stored as `igraph` package (Csárdi and Nepusz 2006; Csárdi et al. 2024) graph objects for further analysis.

How terrain is handled in graph edge cost estimation (i.e., assigning half the cost from the terrain at the vertex at each end of the edge).

`igraph` implements betweenness centrality measures (Freeman 1978; Brandes 2001), which we used to identify vertices and/or edges most likely to be traversed by minimum travel time routes across the terrain.

## 4.3 Betweenness centrality limited by radius.

## 4.4 Impact minimizing networks

Tentative proposal for impact minimizing networks based on minimum spanning trees, but noting the issue with respect to directed graphs when these would more correctly be arborescences (Korte and Vygen 2018).

# 5 Discussion

Blah

# 6 Conclusions

Blah

Brandes, U. 2001. A faster algorithm for betweenness centrality. *The Journal of Mathematical Sociology* 25(2). Routledge: 163–177. doi: [10.1080/0022250X.2001.9990249](https://doi.org/10.1080/0022250X.2001.9990249).

Campbell, M. J., P. E. Dennison, B. W. Butler, and W. G. Page. 2019. Using crowdsourced fitness tracker data to model the relationship between slope and travel rates. *Applied Geography* 106: 93–107. doi: [10.1016/j.apgeog.2019.03.008](https://doi.org/10.1016/j.apgeog.2019.03.008).

- Cox, S. C., B. Smith Lyttle, S. Elkind, C. S. Smith Siddoway, P. Morin, G. Capponi, T. Abu-Alam, M. Ballinger, L. Bamber, B. Kitchener, L. Lelli, J. F. Mawson, A. Millikin, N. Dal Seno, L. Whitburn, T. White, A. Burton-Johnson, L. Crispini, D. Elliot, S. Elvevold, J. W. Goodge, J. A. Halpin, J. Jacobs, A. P. Martin, E. Mikhalsky, F. Morgan, P. Scadden, J. Smellie, and G. S. Wilson. 2023a. A continent-wide detailed geological map dataset of Antarctica. *Scientific Data* 10(1): 250. doi: [10.1038/s41597-023-02152-9](https://doi.org/10.1038/s41597-023-02152-9).
- Cox, S. C., B. Smith Lyttle, S. Elkind, C. S. Smith Siddoway, P. Morin, G. Capponi, T. Abu-Alam, M. Ballinger, L. Bamber, B. Kitchener, L. Lelli, J. F. Mawson, A. Millikin, N. Dal Seno, L. Whitburn, T. White, A. Burton-Johnson, L. Crispini, D. Elliot, S. Elvevold, J. W. Goodge, J. A. Halpin, J. Jacobs, E. Mikhalsky, A. P. Martin, F. Morgan, J. Smellie, P. Scadden, and G. S. Wilson. 2023b. The GeoMAP (v.2022-08) continent-wide detailed geological dataset of Antarctica. PANGAEA. doi: [10.1594/PANGAEA.951482](https://doi.org/10.1594/PANGAEA.951482).
- Csárdi, G., and T. Nepusz. 2006. The igraph software package for complex network research. *Interjournal Complex Systems* 1695: 1–9.
- Csárdi, G., T. Nepusz, V. Traag, S. Horvát, F. Zanini, D. Noom, and K. Müller. 2024. *igraph: Network analysis and visualization in R*. Manual. doi: [10.5281/zenodo.7682609](https://doi.org/10.5281/zenodo.7682609).
- Deutscher und Österreichischer Alpenverein. 2011. *Wegehandbuch des Alpenvereins*. Deutscher Alpenverein.
- Etherington, T. R. 2012. Least-cost modelling on irregular landscape graphs. *Landscape Ecology* 27: 957–968. doi: [10.1007/s10980-012-9747-y](https://doi.org/10.1007/s10980-012-9747-y).
- Felden, J., L. Möller, U. Schindler, R. Huber, S. Schumacher, R. Koppe, M. Diepenbroek, and F. O. Glöckner. 2023. PANGAEA - Data Publisher for Earth & Environmental Science. *Scientific Data* 10(1): 347. doi: [10.1038/s41597-023-02269-x](https://doi.org/10.1038/s41597-023-02269-x).
- Freeman, L. C. 1978. Centrality in social networks conceptual clarification. *Social Networks* 1(3): 215–239. doi: [10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7).
- Goodchild, M. F. 2020. Beyond Tobler’s Hiking Function. *Geographical Analysis* 52(4): 558–569. doi: [10.1111/gean.12253](https://doi.org/10.1111/gean.12253).
- Herzog, I. 2010. Theory and practice of cost functions. In *CAA’2010 Fusion*

- of Cultures: *Proceedings of the 38th Annual Conference on Computer Applications and Quantitative Methods in Archaeology*, ed. F. Contreras, M. Farjas, and F. J. Melero, 375–382. Granada, Spain: BAR Publishing.
- Herzog, I. 2020. Spatial analysis based on cost functions. In *Archaeological Spatial Analysis: A Methodological Guide*, ed. M. Gillings, P. Hacigüzeller, and G. Lock, 333–358. Routledge.
- Imhof, E. 1950. *Gelände und Karte*. Zurich: Eugen Rentsch Verlag.
- Irmischer, I. J., and K. C. Clarke. 2018. Measuring and modeling the speed of human navigation. *Cartography and Geographic Information Science* 45(2). Taylor & Francis: 177–186. doi: [10.1080/15230406.2017.1292150](https://doi.org/10.1080/15230406.2017.1292150).
- Korte, B., and J. Vygen. 2018. Spanning Trees and Arborescences. In *Combinatorial Optimization: Theory and Algorithms*, ed. B. Korte and J. Vygen, 133–157. Berlin, Heidelberg: Springer. doi: [10.1007/978-3-662-56039-6\\_6](https://doi.org/10.1007/978-3-662-56039-6_6).
- Langmuir, E. 1984. *Mountaineering and Leadership: A Handbook for Mountaineers and Hillwalking Leaders in the British Isles*. Edinburgh: Scottish Sports Council.
- Llobera, M., and T. J. Sluckin. 2007. Zigzagging: Theoretical insights on climbing strategies. *Journal of Theoretical Biology* 249(2): 206–217. doi: [10.1016/j.jtbi.2007.07.020](https://doi.org/10.1016/j.jtbi.2007.07.020).
- Márquez-Pérez, J., I. Vallejo-Villalta, and J. I. Álvarez-Francoso. 2017. Estimated travel time for walking trails in natural areas. *Geografisk Tidsskrift-Danish Journal of Geography* 117(1). Routledge: 53–62. doi: [10.1080/00167223.2017.1316212](https://doi.org/10.1080/00167223.2017.1316212).
- Martínez, D., L. Velho, and P. C. Carvalho. 2005. Computing geodesics on triangular meshes. *Computers & Graphics* 29(5): 667–675. doi: [10.1016/j.cag.2005.08.003](https://doi.org/10.1016/j.cag.2005.08.003).
- Naismith, W. W. 1892. Excursions: Cruach Ardran, Stobinian, and Ben More. *Scottish Mountaineering Club Journal* 2(3): 133.
- Perna, A., and T. Latty. 2014. Animal transportation networks. *Journal of The Royal Society Interface* 11(100). Royal Society: 20140334. doi: [10.1098/rsif.2014.0334](https://doi.org/10.1098/rsif.2014.0334).
- Prisner, E., and P. Sui. 2023. Hiking-time formulas: A review. *Cartography and Geographic Information Science* 50(4): 421–432. doi:

[10.1080/15230406.2023.2197625](https://doi.org/10.1080/15230406.2023.2197625).

- R Core Team. 2024. *R: A language and environment for statistical computing*. Manual. Vienna, Austria: R Foundation for Statistical Computing.
- Rees, W. G. 2004. Least-cost paths in mountainous terrain. *Computers & Geosciences* 30(3): 203–209. doi: [10.1016/j.cageo.2003.11.001](https://doi.org/10.1016/j.cageo.2003.11.001).
- Tobler, W. R. 1993. *Three Presentations on Geographical Analysis and Modeling: Non- Isotropic Geographic Modeling; Speculations on the Geometry of Geography; and Global Spatial Analysis*. Technical {{Report}} 93-1. Santa Barbara, CA: National Center for Geographic Information and Analysis.
- Verhagen, P., L. Nuninger, and M. R. Groenhuijzen. 2019. Modelling of pathways and movement networks in archaeology: An overview of current approaches. In *Finding the Limits of the Limes*, ed. P. Verhagen, J. Joyce, and M. R. Groenhuijzen, 217–249. Cham: Springer International Publishing. doi: [10.1007/978-3-030-04576-0\\_11](https://doi.org/10.1007/978-3-030-04576-0_11).
- Winkler, K., H.-P. Brehm, and J. Haltmeier. 2010. *Bergsport Sommer: Technik, Taktik, Sicherheit*. SAC-Verlag.
- Wood, A., W. Mackaness, T. I. Simpson, and J. D. Armstrong. 2023. Improved prediction of hiking speeds using a data driven approach. Edited by Yuxia Wang. *PLOS ONE* 18(12): e0295848. doi: [10.1371/journal.pone.0295848](https://doi.org/10.1371/journal.pone.0295848).