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 betweennness 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 Östereichischer 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 θ 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	$ae^{b\lvert\theta-c\rvert}$	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 parameterisation of the functions in Figure 1 have been chosen for illustrative purposes only, although the parameter values for the Exponential hiking function shown are $v = 6e^{-3.5\lvert\theta+0.05\rvert}$ as suggested in Tobler (1993). These parameter values for the exponential functional form 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 function 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

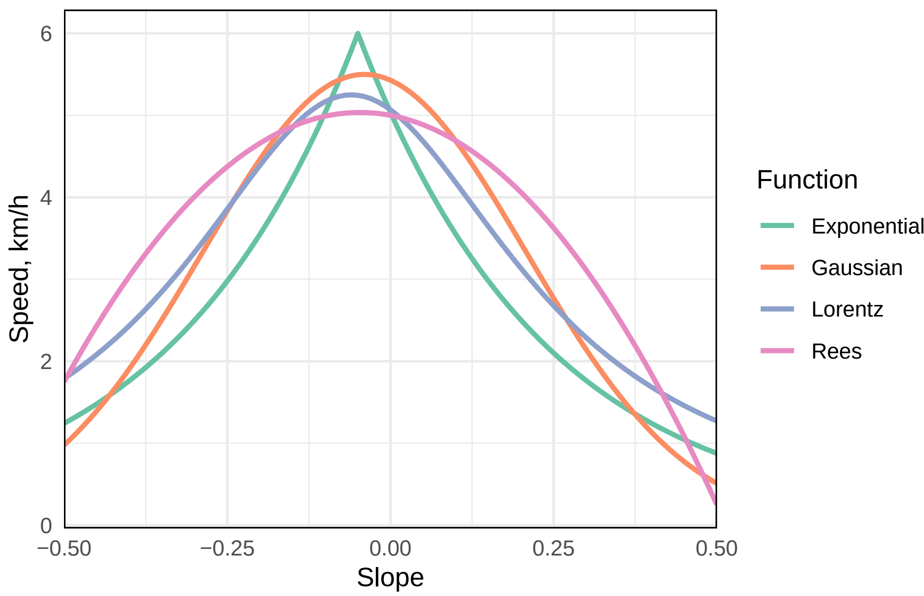

Figure 1: Example hiking functions.

more reliably generate zig-zag or 'switchback' movement behaviour on steep slopes (Llobera and Sluckin 2007). However, these approaches are hard to implement, and empirically-derived and locally-specific hiking functions as reported by Márquez-Pérez et al. (2017), Irmischer and Clarke (2018), and Campbell et al. (2019) are the approach we have taken in the present work, 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). The hiking function reported 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 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 has the additional advantage of allowing for distinctly different effects of slope on speed. Details of this approach are reported and discussed in Section 4.1.

2.2 Representing the landscape

Terrain is asymmetric with respect to movement, therefore: landscapes as graphs

Edge costs in graphs modelled based on slope and a hiking function

Movement cost for graph edges derived from hiking functions

Graph vertices augmented with landcover information (in this context limited to surface geology)

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 scatterplot/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

Assignment of nodes as a lattice of equally spaced locations. Choice of rectangular/square grid vs. hexagonal. Square grids are easily generated and also offer some computational advantages if, for example, the grid is aligned with an underlying digital elevation model (DEM). However, square grids present the problem of choosing between two different potential local neighbourhoods, based on orthogonal adjacency only, or on orthogonal and diagonal adjacency (respectively referred to as rook and queen's case adjacency, or as von Neumann and Moore neighbourhoods, in the spatial analysis and cellular automaton literatures). Excluding diagonal adjacency yields implausible diamond shaped isochrones on level terrain. Including diagonal adjacency can lead to seemingly contradictory results when alternate corners of a square are at different heights.

Given these challenges we adopt a hexagonally arranged triangular lattice. This avoids any difficulties with varying distances (i.e. network edge lengths) between nodes. Note that this is necessarily an approximation, and on extensive flat terrains will yield unrealistic hexagonal isochrones. It also presents some computational disadvantages in assigning elevations to graph nodes where interpolation of heights is required due to mismatches between hexagon centres and cells in grid-based DEMs.

[FIGURE showing square and hexagonal lattices and issues]

Alternative approaches might estimate movement based on generating multiple random approximately evenly spaced nodes across a terrain and combining the resulting estimated travel times (see Etherington 2012), or deploying methods for estimating geodesics on complex triangular meshes (Martínez et al. 2005) but both approaches are more computationally demanding than the current application justifies. In any case, the latter method does not allow for asymmetry in traversal times as arising due to slopes.

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.

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.

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.

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.

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.

Deutscher und Östereichischer 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.

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.

Freeman, L. C. 1978. Centrality in social networks conceptual clarification. *Social Networks* 1(3): 215–239. doi: 10.1016/0378-8733(78)90021-7.

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.

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.

Langmuir, E. 1984. *Mountaincraft 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.

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.

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.

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.

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.

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.

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.

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.