

Using crowdsourced fitness tracker data to model the relationship between slope and travel rates

Michael J. Campbell^{a,*}, Philip E. Dennison^b, Bret W. Butler^c, Wesley G. Page^c

^a Department of Geosciences, Fort Lewis College, 1000 Rim Drive, Durango, CO, 81301, United States

^b Department of Geography, University of Utah, 332 South 1400 East, Salt Lake City, UT, 84112, United States

^c USDA Forest Service, Rocky Mountain Research Station, Missoula Fire Sciences Laboratory, 5775 Highway 10 West, Missoula, MT, 59808, United States

ARTICLE INFO

Keywords:

Fitness tracker
GPS data
Travel rate
Hiking
Running
Slope
Least cost path

ABSTRACT

One of the critical factors affecting travel rates while hiking, jogging, or running along a trail is the slope of the underlying terrain. Models for predicting this effect have been used in a wide variety of scientific and applied contexts, including recreation planning, search and rescue, wildland firefighter safety, social network analysis, and recreating historical human movement patterns. Despite their wide use, these models are based on datasets with very small sample sizes that were collected without using instantaneous measures of travel rate and assume symmetrical effects about the slope of maximum travel rate. These models also typically resulted in a single mathematical function, ignoring the significant variability that can occur between a fast and a slow individual, or between walking and running travel rates. In this study we modeled travel rates using a database of GPS tracks from 29,928 individuals representing 421,247 individual hikes, jogs, and runs on trails in and around Salt Lake City, Utah for an entire year between July 1, 2016 and June 30, 2017. Three widely-used probability distribution functions (Laplace, Gauss, and Lorentz) were used to predict travel rates based on terrain slope along segments of trails with uniform slopes. To account for the variability in travel rates between fast and slow movement, a series of travel rate models were generated to predict travel rate percentiles, ranging from the 1st to the 99th, thus providing a flexible basis for predicting travel rates as a function of slope. The large number of samples allowed us to introduce a novel term that accounts for asymmetry in travel rates on uphill and downhill slopes. All three functions performed well, with Lorentz percentile models averaging an R^2 of 0.958 and a mean absolute error (MAE) of 0.078 m/s, Laplace with R^2 of 0.953 and MAE of 0.088 m/s, and Gauss with R^2 of 0.949 and MAE of 0.090 m/s. All three functions performed notably better at estimating lower travel rate percentiles (e.g. 5th: $R^2_{\text{Lorentz}} = 0.941$; $R^2_{\text{Laplace}} = 0.940$; $R^2_{\text{Gauss}} = 0.934$) as compared to higher (e.g. 95th: $R^2_{\text{Lorentz}} = 0.914$; $R^2_{\text{Laplace}} = 0.913$; $R^2_{\text{Gauss}} = 0.908$), indicating greater consistency in walking rates than the fastest running rates. Lorentz outperformed the other functions for the widest range of percentiles (5th, 30th–90th), and thus is recommended for use as a flexible travel rate prediction function. However, Laplace tended to produce the best results at moderately-low travel rate percentiles (10th–25th), suggesting a combination of the two models could produce the highest accuracies. The results of this research provide a sound basis for future studies aiming to estimate travel rates while hiking or running along slopes.

1. Introduction

There are three general relationships that apparently govern the effect of terrain slope on pedestrian movement: (1) higher energy output is required with increasing (steeper) uphill slope; (2) braking against gravity with increasing downhill slope and ensuring the maintenance of safety through careful foot placement; and (3) as a result of (1) and (2), with increasing slope in both directions, one tends to move more slowly, though these effects are not symmetrical about 0° in slope

(Minetti, Moia, Roi, Susta, & Ferretti, 2002). Although these relationships are fairly intuitive and innate to the human bipedal experience, the quantification of these effects is quite complex, due to the variability in modes of travel (e.g. walking vs. sprinting), distances of travel (e.g. 100-m dash vs. marathon), and between the physical characteristics of individuals (e.g. variation in stride length). Despite this inherent complexity, several mathematical functions aimed at quantifying the effect of slope on walking and/or hiking travel rates have been developed, with early efforts spanning as far back as the 19th century

* Corresponding author.

E-mail address: mcampbell@fortlewis.edu (M.J. Campbell).

<https://doi.org/10.1016/j.apgeog.2019.03.008>

Received 17 July 2018; Received in revised form 6 February 2019; Accepted 16 March 2019

0143-6228/ © 2019 Elsevier Ltd. All rights reserved.

(Campbell, Dennison, & Butler, 2017; Davey, Hayes, & Norman, 1994; Irmischer & Clarke, 2018; Naismith, 1892; Rees, 2004; Tobler, 1993, p. 24). These functions broadly show that steeper slopes result in slower movement (Fig. 1). Differences between these previously proposed relationships likely result from the varied scientific approaches taken to derive the mathematical functions, including the use of primary vs. secondary travel rate data, the techniques used for tracking/timing study subjects, the travel distances used as a baseline for experimentation, the statistical function used to fit the data, the assumption of symmetry about the 0° slope value, and, perhaps most importantly, the sample size of study subjects (Table 1).

With the recent proliferation of GPS-enabled smartphones, we now have the capability of collecting large quantities of travel rate information in real time from a diverse population of users. Popular fitness tracking applications, such as Strava, allow users to collect, share, and compare travel rates along roads and trails while hiking, running, and biking. Rather than relying on sparse travel rate information from relatively few individuals, we can leverage the power of a vast database of GPS tracks to derive a much more robust and broadly-applicable mathematical model for estimating the effect of slope on travel rates. In addition, all of the aforementioned functions are based on the prediction of travel rates while walking or hiking along a slope. To the authors' knowledge, no models for the prediction of running travel rates as a function of slope currently exist. This could be due to the greater variability between individuals' running travel rates as compared to their walking travel rates and the associated difficulty of developing a generalizable predictive model. With the use of a fitness tracking database, where modes of pedestrian travel range from walking to jogging to sprinting, such a model could be realized.

In this study, some of the key limitations inherent to existing slope-travel rate models are addressed by employing the use of a very large, crowdsourced database of GPS tracks representing trail hikes, jogs, and

Table 1

Sample sizes of study subjects used as a basis for deriving several existing travel rate functions.

Study	Number of Subjects
Tobler (1993)	Unknown (secondary data)
Naismith (1892)	1
Davey et al. (1994)	2
Rees (2004)	Unknown (based on 10 hikes)
Campbell et al. (2017)	37
Irmischer and Clarke (2018)	200

runs. Specifically, the primary objectives of this study are to: (1) develop a flexible model for estimating travel rates as a function of slope based on a range of travel rate percentiles from the 1st (slow walking/hiking) to the 99th (fast running/sprinting) from a large database of GPS tracks along trails; (2) compare the statistical and geospatial similarities and differences between three popular distribution functions used as a basis for modeling slope-controlled travel rates; and (3) apply the best-fitting function to the simulation of total hiking time along a popular trail and compare the results to existing travel rate functions and a validation dataset.

2. Background

Tobler's hiking function (THF; Tobler, 1993, p. 24) is one of the most widely-used existing models for the estimation of travel rates as a function of slope (Alegana et al., 2012; Contreras, 2011; Doherty, Guo, Doke, & Ferguson, 2014; Fryer, Dennison, & Cova, 2013; Herzog, 2010; Jennings & Craig, 2001; Jensen, 2003; Kantner, 2004; Kunitz, Lagree, & Weinig, 2017; Márquez-Pérez, Vallejo-Villalta, & Álvarez-Francoso, 2017; McCoy et al., 2011; Mink, Ripy, Bailey, & Grossardt, 2009; Richards-Rissetto & Landau, 2014; Ripy et al., 2014; Rogers, Collet, &

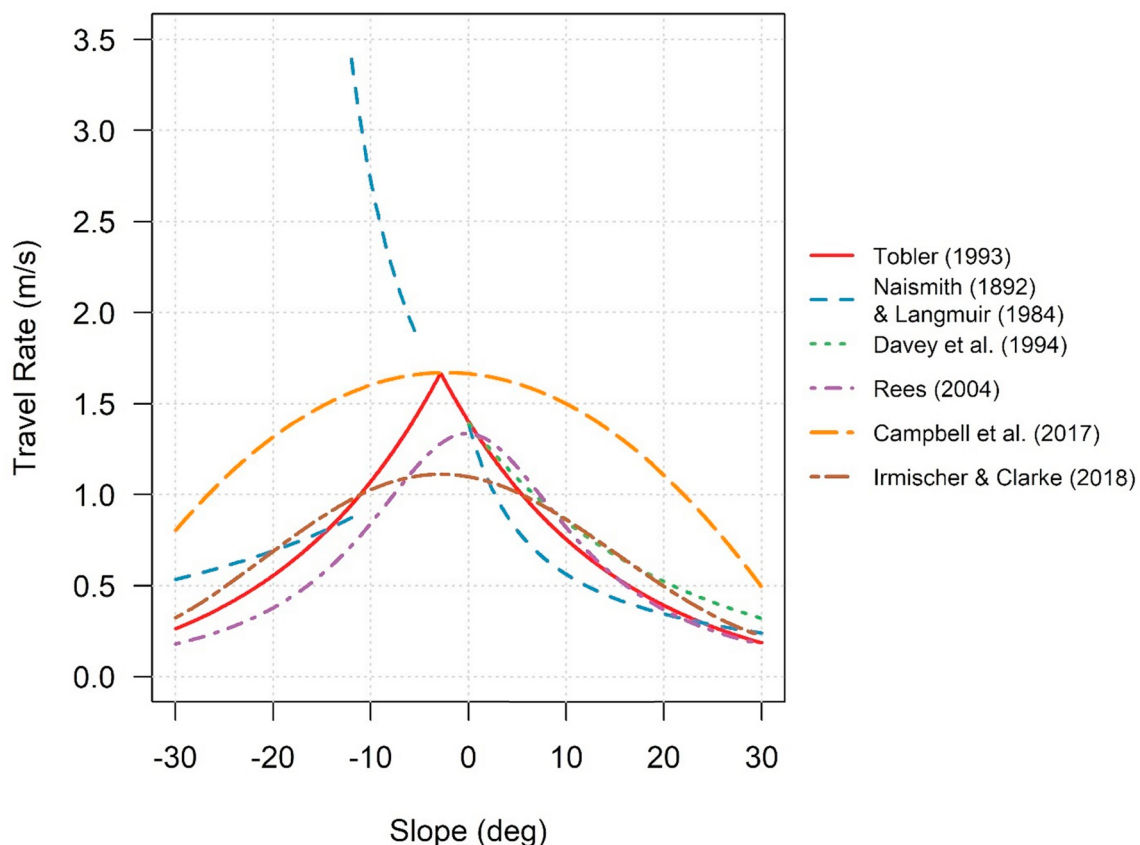


Fig. 1. Existing slope-travel rate functions, where downhill slopes are negative and uphill slopes are positive.

Lugon, 2014; Rogers, Fischer, & Huss, 2014; Taliaferro, Schriever, & Shackley, 2010; Verhagen & Jeneson, 2012; White, 2015; Wood & Schmidtlein, 2012). It is calculated as:

$$v = 1.6 \times \exp^{-3.5 \times |\tan \theta + 0.05|} \quad (1)$$

where v is velocity in m/s and θ is the slope angle in degrees. The function was derived from empirical data presented by Imhof (1950). As Herzog (2010) points out, however, these foundational data from Imhof (1950) were not presented in their raw form; instead, he provided a figure with aggregated slope-based travel time isolines, to which Tobler (1993, p. 24) fit a mathematical function, with no reporting on the statistical model fit or accuracy. Like many travel rate functions, THF takes a double-exponential form, which is characterized by a sharp peak, representing a maximum travel rate, and rapidly-decreasing travel rates on either side of the peak (Fig. 1). While the function is mirrored on both sides of a peak travel rate, THF assumes that the peak travel rate occurs on a slight downhill slope. According to THF, one travels most efficiently along slightly-downhill slopes of -2.86° , with an estimated hiking travel rate of approximately 1.67 m/s.

THF has been used for estimating travel rates in a range of contexts, including the location of missing persons for search and rescue (Doherty et al., 2014), assessment of urban social interaction and integration (Richards-Rissetto & Landau, 2014), simulation of pedestrian evacuation times in the event of a tsunami (Wood & Schmidtlein, 2012), analysis of hiking times and difficulties along trails for recreation planning (Márquez-Pérez et al., 2017), modeling of pedestrian health care facility accessibility in developing nations (Alegana et al., 2012), and estimation of wildland firefighter escape route travel time to safety (Fryer et al., 2013). By far, the discipline that has employed THF the most is that of archaeology (White, 2015), where it has been used to recreate historical migration routes (Jensen, 2003; Kantner, 2004; Kunitz et al., 2017; Verhagen & Jeneson, 2012), analyze trade and geopolitical interactions (Contreras, 2011; Jennings & Craig, 2001; McCoy et al., 2011; Taliaferro et al., 2010), and prediction of archaeological site locations (Mink et al., 2009; Ripy et al., 2014; Rogers, Collet, et al., 2014; Rogers, Fischer, et al., 2014).

Perhaps the second most widely-used slope-travel rate function is the one developed in 1892 by Scottish mountaineer William Naismith (Naismith, 1892). “Naismith’s Rule”, based apparently on personal experience, suggested that one should need one hour for every three horizontal miles traveled, with an additional hour budgeted for every 2000 feet of vertical ascent. Though it was not presented as a mathematical function per se, Naismith’s Rule can be defined as follows:

$$v = \frac{1}{0.72 + 6 \times \tan \theta} \quad (2)$$

Naismith’s Rule is similar in form to THF (Fig. 1). In fact, the travel rates are nearly identical on flat slopes ($v_{Tobler, 0^\circ} = 1.39$ m/s; $v_{Naismith, 0^\circ} = 1.40$ m/s). However, Naismith’s Rule makes no mention of downhill travel, and thus is only used to estimate travel rates on positive slopes. To correct for this omission, Langmuir (1984) suggested modifying Naismith’s rule, such that on steep downhill slopes ($\theta \leq -12^\circ$):

$$v = \frac{1}{0.72 - 2 \times \tan |\theta|} \quad (3)$$

and on less steep downhill slopes ($-12^\circ < \theta \leq -5^\circ$):

$$v = \frac{1}{0.72 + 2 \times \tan |\theta|} \quad (4)$$

Although there is no specific mention of slightly-downhill slopes ($-5^\circ < \theta < 0^\circ$), Irmischer and Clarke (2018) point out that the popular open-source GIS platform GRASS’ *r.walk* function, which is based on the Langmuir corrections for Naismith’s Rule, simply assumes that the travel rate on these slopes is constant and is the same as the travel rate on flat slopes (1.40 m/s). The Langmuir corrections result in

a very peculiar and counter-intuitive form (Fig. 1), suggesting that one travels fastest at a downhill slope of -12° (3.39 m/s), but then at -12.1° , the travel rate would drop precipitously to (0.87 m/s) – a highly improbable set of circumstances. Naismith’s Rule, with and without the Langmuir corrections, has been applied frequently, though seemingly not as often as THF. It has been used to assess the remoteness of wild lands in Scotland (Carver, Comber, McMorran, & Nutter, 2012), analyze trail usage and travel time information for recreation planning purposes (Chiou, Tsai, & Leung, 2010), study the perceptions of safety while moving through an urban environment (De Silva, Warusavitharana, & Ratnayake, 2017), develop a sampling design for forest inventory (Tomppo et al., 2014) as well as several archaeological applications, including social network analysis (Gravel-Miguel, 2016), historical route recreation (Tomppo et al., 2014), and archaeological site accessibility assessment (Henry, Belmaker, & Bergin, 2017).

In addition to these two popular slope-travel rate functions, several others, both more recent and less widely-cited, have been developed. Davey et al. (1994) presented a function based on a two-subject treadmill experiment, which allowed for an adjustment factor to fit the slope-travel rate function based on a baseline individual velocity that is sustainable over an unspecified long distance on flat ground (v_0), as follows:

$$v = v_0 \times \exp^{-0.049 \times \theta} \quad (5)$$

Like Naismith’s Rule, the function presented by (Davey et al., 1994) does not have a downhill component; however, the adjustment factor (v_0) provides a more flexible framework for travel rate estimation, given differences in fitness and endurance levels. This function, adjusted for a baseline rate of $v_0 = 1.40$ m/s can be seen in Fig. 1.

Rees (2004) presented a polynomial travel rate model based on the compiled results of 10 individual GPS-tracked walks that, unlike other functions, did not come to a sharp, peak travel rate, and instead took a more bell curve-like form (Fig. 1). The resulting equations is as follows:

$$v = \frac{1}{0.75 + 0.09 \times \tan \theta + 14.6 \times (\tan \theta)^2} \quad (6)$$

Campbell et al. (2017) performed an experiment with 37 subjects to determine the effects of slope, vegetation density, and ground surface roughness on travel rates in an off-trail environment. Like (Rees, 2004), the resulting function does not feature a strong peak travel rate; unlike (Rees, 2004), the effects are not symmetrical about zero slope, with the peak travel rate (1.67 m/s) occurring at -2.3° , similar to THF (Fig. 1). Although the resulting function contained terms for the latter two landscape conditions, if no vegetation and a smooth surface are assumed, the equation for estimating travel rates based on slope alone becomes:

$$v = 1.662 - (5.191 \times 10^{-3}) \times \theta - (1.127 \times 10^{-3}) \times \theta^2 \quad (7)$$

Most recently, Irmischer and Clarke (2018) presented a slope-travel rate function based on the real-time GPS tracking of 200 United States Military Academy cadets as they traversed varied terrain during training exercises. They compared the resulting travel rates to both THF and the Langmuir-corrected Naismith Rule, finding that neither fit their data very well, particularly given the lack of a strong peak in travel rates. They found that their data closely aligned with a Gauss curve (Fig. 1), resulting in the following function:

$$v = 0.11 + \exp^{\frac{-(100 \times \tan \theta + 5)^2}{1800}} \quad (8)$$

As Fig. 1 highlights, with the exception of the Langmuir corrections, the functions all fall within a fairly consistent range of travel rates. Differences between the functions likely stem from the differences in sample size (Table 1) and sample populations, and differences in analytical approaches ranging from interpretation of existing empirical data (THF) to timed treadmill and field experiments (Davey et al., 1994 and Campbell et al., 2017, respectively), and GPS-tracked field

experiments (Irmischer & Clarke, 2018).

Mathematical functions for predicting travel rates are useful in their own right; however, they are often made more practical through their application in a geospatial context (e.g. Campbell et al., 2017; Contreras, 2011; Doherty et al., 2014; Verhagen & Jeneson, 2012; Wood & Schmidtlein, 2012). When used in conjunction with digital elevation model (DEM) data, these slope-travel rate functions allow for the computation of travel cost, which represents the time (e.g. seconds) it would take a pedestrian to traverse a given raster pixel on the ground, according to that pixel's slope and the direction of movement. Provided that the slope-travel rate function is accurate, this allows for the estimation of travel time between any two (or more) locations along a least-cost path. To assume, however, that slope is the only impedance to pedestrian mobility, is an oversimplification of reality, where other landscape conditions such as the ground surface condition and the presence of vegetation can limit efficient movement (Alexander, Baxter, & Dakin, 2005; Butler, Cohen, Putnam, Bartlette, & Bradshaw, 2000; Campbell et al., 2017), as well as impassible features such as water bodies. Of the models presented, only Campbell et al. (2017) attempt to account for ground surface condition.

3. Methods

The travel rate data used in this study were obtained from Strava, Inc. (San Francisco, California, USA) via their Strava Metro program. Strava is a popular fitness tracking and social networking app that allows users to track their movement while hiking, running, and cycling using GPS and compare their travel rates to their peers. As such, Strava's databases contain perhaps the most comprehensive record of travel rate information in existence, with 136 million runs tracked in 2017 spanning more than 700 million miles worldwide (Strava, 2018). The company aggregates and anonymizes the data and makes them available to planning organizations and researchers through the Strava Metro program. Strava Metro data have been used by researchers to study urban cycling patterns (Musakwa & Selala, 2016), air pollution exposure while cycling (Sun & Mobasheri, 2017), the relative effectiveness of improving cycling infrastructure such as adding bike lanes (Heesch & Langdon, 2017), and many other cycling-specific applications, though the pedestrian data (hiking, jogging, running) have received relatively scarce attention in the scientific literature.

To obtain their pedestrian travel data, Strava Metro requested that we provide them with a spatial dataset containing individual line features representing segments of trails, to which their GPS track data was aggregated using their proprietary algorithms. The process for creating those trail segments proceeded as follows. We first selected the trails in and around Salt Lake City, Utah as a basis for study (Fig. 2), given the abundance of trails, the large population, diversity of terrain, and presence of existing lidar data for use in deriving a high spatial resolution terrain model. The study area in Fig. 2 is defined by the extent of a lidar data collection that took place in 2006–2007. It is a 3888 km² area centered at approximately 40° 43' 32" N, 111° 51' 40" W, spanning a 7-county area in northern-central Utah, with the majority of the area falling within Salt Lake County. As of 2010, there were approximately 1.36 million people living within the study area extent, according to Census Block data (United States Census Bureau, 2010). It lies at the interface between the Basin and Range and Middle Rocky Mountain geologic provinces (Utah Geological Survey, 2000), featuring a diversity of terrain and geology, with a wide range of elevations, from 1237 m (4058 ft) on the lakebed sediment-deposited basin floor to 3502 m (11490 ft) in the glacially-carved range peaks.

Lidar and trails data were obtained from the Utah Automated Geographic Reference Center (AGRC, <https://gis.utah.gov/>). The lidar data were collected between 2006 and 2007 by Horizons, Inc. on behalf of the State of Utah with an average post spacing of 2 m and a vertical root mean square position error of approximately ± 5 cm. The raw lidar data were processed by AGRC in order to derive a raster digital terrain

model at a spatial resolution of 2 m, which was the basis of the terrain data used in our study. The statewide trails dataset was clipped to the extent of the lidar data.

Slope varies continuously at a finer spatial scale than the travel rate data, so we developed an algorithm inspired by the LandTrendr temporal segmentation process (Kennedy, Yang, & Cohen, 2010) to split trails into segments of relatively homogeneous slope (Fig. 3). The algorithm begins by placing points along every trail at a 2 m interval and extracts the lidar-derived terrain elevation at each point. It then labels the first and last points along the trail as inflection points and creates a trail segment by connecting them. The root mean squared difference between the trail segment and the true surface is then calculated. If this value exceeds a pre-defined threshold (1 m), an additional inflection point is placed on the single point that has the largest difference from the trail segment. This new inflection point is used to split the trail into two new segments. This process is repeated until all of the trail segments are below the root mean squared difference threshold.

Trails were additionally split at each intersection with other trails, where such intersections occurred. All trail segments under 50 m in length were eliminated to minimize the potential travel rate noise that would likely emerge from such small distances. Trail segment slope was calculated as the arctangent of each trail segment's elevation gain divided by the horizontal segment distance. Then, slope distance, a more realistic measure of true distance traveled, was calculated as the horizontal distance of each segment divided by the cosine of the slope. In all, 9870 trail segments from 1233 individual trails, totaling 1409.68 km in distance were delivered to Strava (Table 2), with which they were able to perform their internal GPS database querying, extracting, and aggregating procedure.

In return, Strava provided data from 29,928 anonymous individuals (Table 3) representing 421,247 individual hikes, jogs, and runs on trails in and around Salt Lake City, Utah for an entire year between July, 2016 and June, 2017. The database we received contained the attributes seen in Table 4. The median total distance traveled per activity was 6544 m (4.07 miles), and the median total time per activity was approximately 49 min. After splitting these activities up by the trail segments, there were 2.99 million individual travel rate records. Several steps were taken to reduce potential sources of bias in the data: (1) remove data from all trails marked as paved or improved to focus solely on natural surface trails; (2) remove data from slopes greater than 30°, given their sparseness; (3) remove data where multiple individuals were exercising together to focus on individual travel rates; (4) remove data flagged as being a part of a commute to limit focus to intentional fitness activities; and (5) remove all travel rates above 10 m/s, as this value approximates the limits of human performance and travel rates above that threshold are believed to be due to a variety of factors including GPS error (Ranacher, Brunauer, Trutschnig, Spek, & Reich, 2016), aggregation error, and/or user error (e.g. forgetting to stop GPS tracking after getting into a car, tagging the wrong fitness activity, etc.). As a result, 1.05 million records remained, forming the basis of the subsequent analyses.

Given the wide variety of travel rates that emerged across the range of slopes analyzed ($\pm 30^\circ$), it became quite clear that no single travel rate model accounts for sufficient variability to be useful in a broad range of applications. In the raw Strava data, there was no ability to control for whether someone was hiking, jogging, or running, nor was there any individual-level information to distinguish between fit or unfit individuals, males or females, young or old. Accordingly, we first calculated a series of per-slope-degree travel rate percentiles (1st, 5th – 95th, by 5, and 99th) (Fig. 4).

Given the visual patterns that emerged from the percentile analysis, we sought to determine the best mathematical functions that could closely fit these percentiles. Without associated heart rate or energy consumption information, there is no way to link these percentiles to modes of travel (e.g. walking, jogging, sprinting), but they provide an estimate of the range of possible travel rates, from slow (1st) to fast

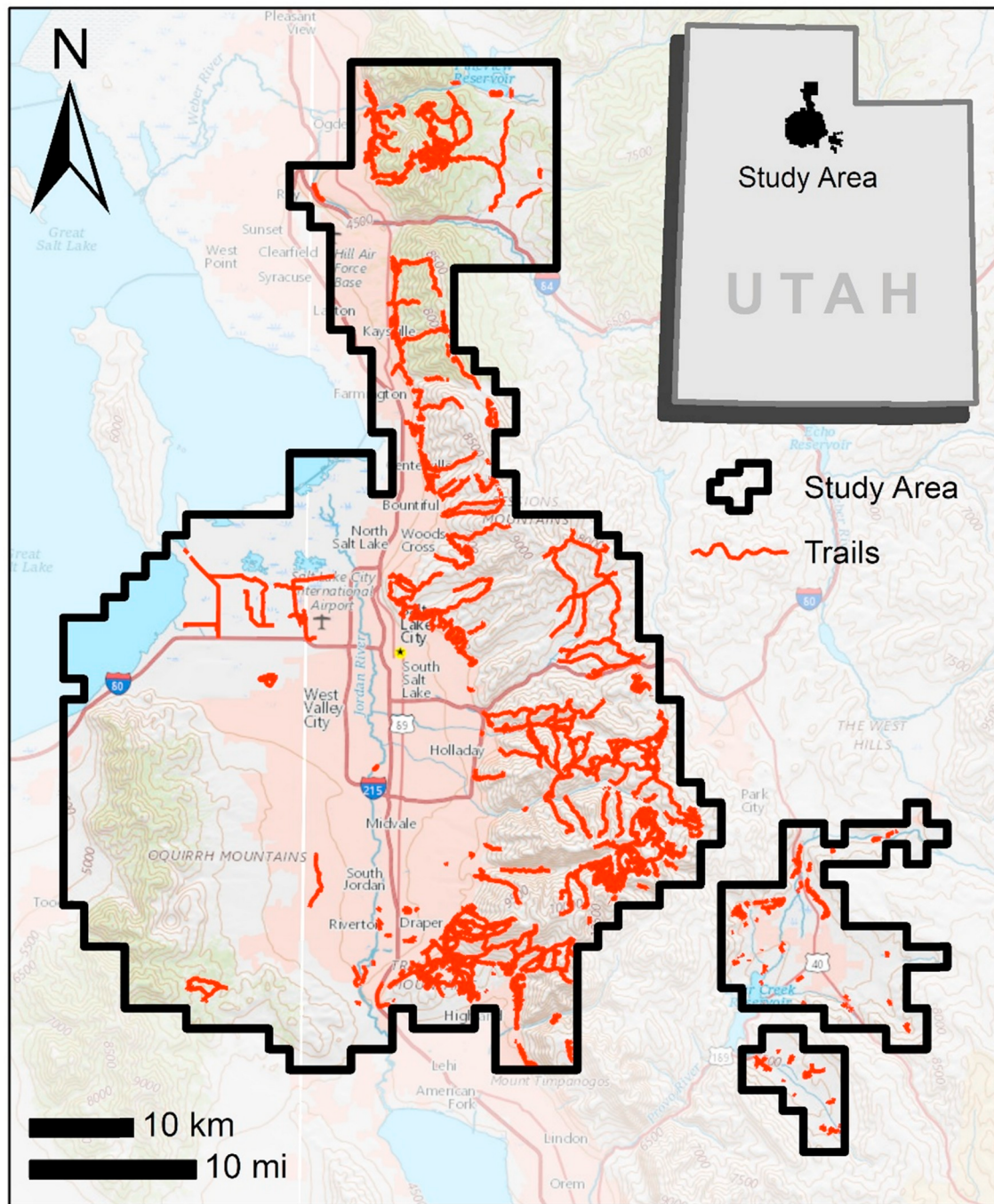


Fig. 2. Study area map highlighting all non-paved trail segments longer than 50 m with slopes < 30° used as a basis for slope-travel rate analysis.

(99th) movement. To fit these percentiles, we tested three common probability distribution functions: Laplace (also known as double exponential), Gauss (also known as normal), and Lorentz (also known as Cauchy). These three functions are defined, respectively, below:

$$y_{\text{Laplace}} = \frac{1}{2b} \exp \left(-\frac{|x-a|}{b} \right) \quad (9)$$

$$y_{\text{Gauss}} = \frac{1}{\sqrt{2\pi}b} \exp \left(-\frac{(x-a)^2}{2b^2} \right) \quad (10)$$

$$y_{\text{Lorentz}} = \frac{1}{\pi b \left[1 + \left(\frac{x-a}{b} \right)^2 \right]} \quad (11)$$

where a represents a curve-centering term (akin to the mean of the distribution), b represents a curve-width-widening term (akin to the standard deviation of the distribution). The effects of these parameters can be seen in Fig. 5. The Laplace function comes to a pointed peak value, similar to the travel rate functions of Tobler (1993, p. 24), Naismith (1892), and Davey et al. (1994), whereas the Gauss and Lorentz functions come to a rounded peak value, as in Rees (2004), Campbell et al. (2017), and Irmischer and Clarke (2018). The key difference between the Gauss and Lorentz distributions is the fact that the Gauss distribution has a more rounded peak and flatter tails, whereas the Lorentz distribution has a less rounded peak and less flattened tails.

As with all probability distribution functions, the y values are bounded by [0,1]. Given that the travel rate data are bounded by [0,10], an additional, multiplicative term (c) needs to be included in the

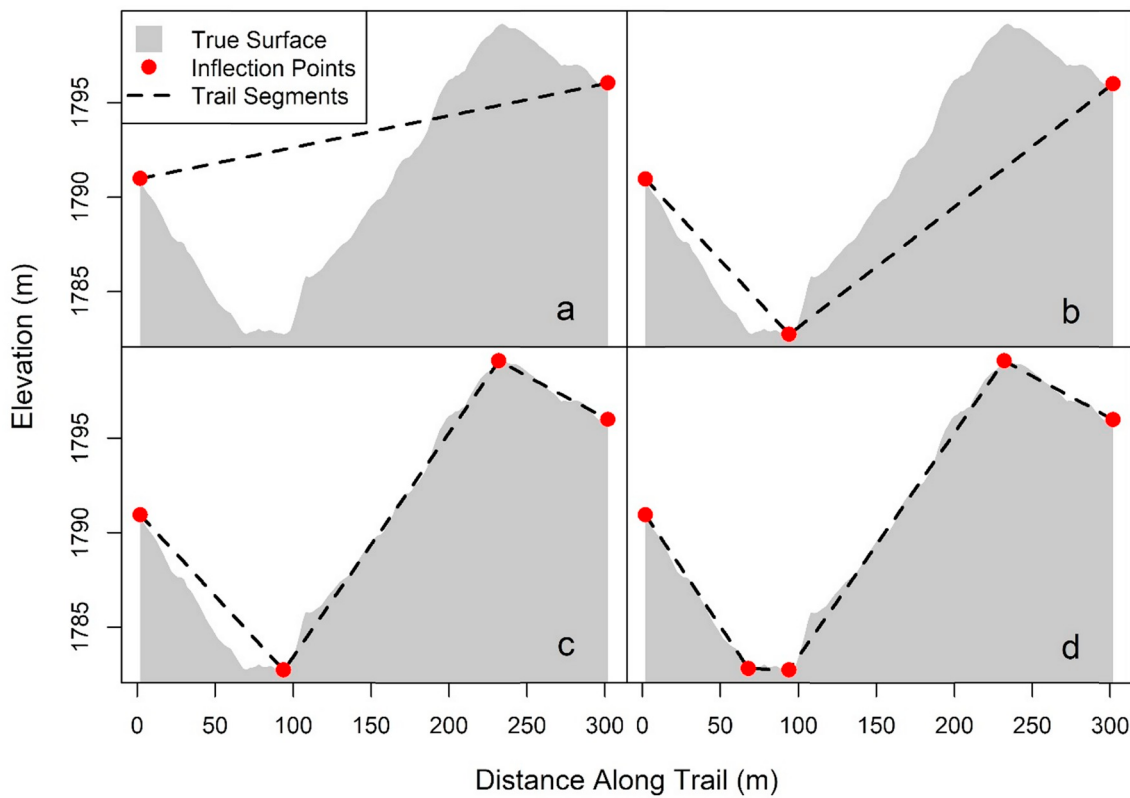


Fig. 3. Trail segmentation technique example, with each iteration (a, b, c, and d) adding a split point to reduce root mean squared difference between the segments and actual trail elevation.

functions. As seen in Fig. 4, none of the travel rate percentiles reach zero; thus, there needs to be an additive term (d) in the functions as well. Lastly, upon close examination of the percentile lines in Fig. 4, it becomes clear that the upslope effects and downslope effects differ, with upslope travel rates being both more consistent among the percentiles and generally lower than the downslope travel rates. Accordingly, a final slope-dependent multiplicative term (e) needs to be added to the functions, allowing for anisotropy in uphill and downhill travel rates that has not been accounted for in previously proposed functions. Substituting the slope (θ ; assumed to be in degrees) for x and the travel rate (v ; assumed to be in m/s) for y , the functions for Laplace, Gauss, and Lorentz travel rate estimation are as follows:

$$v_{Laplace} = c \left(\frac{1}{2b} \exp^{-\frac{|\theta-a|}{b}} \right) + d + e\theta \quad (12)$$

$$v_{Gauss} = c \left(\frac{1}{\sqrt{2\pi}b^2} \exp^{-\frac{(\theta-a)^2}{2b^2}} \right) + d + e\theta \quad (13)$$

Table 3

Age and gender characteristics of anonymous study subjects from Strava Metro database.

Age	Gender		
	Male	Female	Unknown
< 25	1438	861	
25–34	4151	2349	
35–44	5406	2178	
45–54	2869	867	
55–64	914	319	
65–74	182	64	
75–84	16	9	
85–94	2	2	
≥ 95	11	4	
Unknown	4278	2194	1814
Total	19267	8847	1814

Table 2

Summary characteristics of trail segment data delivered to Strava Metro, with total trail segment length (in km) by slope and elevation classes.

Slope (°)	Elevation (m)					Total
	1000–1500	1500–2000	2000–2500	2500–3000	3000–3500	
< 5	313.06	335.67	96.33	105.79	12.80	863.65
5–10	21.76	139.91	88.19	97.35	16.08	363.30
10–15	2.75	41.14	38.20	37.60	8.40	128.09
15–20	0.16	9.11	13.39	12.38	2.44	37.47
20–25	0.13	2.37	4.72	2.97	1.70	11.89
25–30	0.00	0.80	1.16	0.91	0.59	3.45
≥ 30	0.00	0.53	0.30	0.00	0.39	1.22
Total	337.87	529.52	242.28	257.00	42.40	1409.68

Table 4
Strava Metro database field names and descriptions.

Field Name	Description
<i>edge_id</i>	Unique identifier for trail segments for linking back to spatial data
<i>year</i>	Year that the activity was recorded
<i>day</i>	Julian day that the activity was recorded
<i>hour</i>	Hour that the activity was recorded
<i>minute</i>	Minute that the activity was recorded
<i>athlete_count</i>	Total number of people that recorded an activity for a given trail segment at a given time in the forward direction
<i>rev_athlete_count</i>	Total number of people that recorded an activity for a given trail segment at a given time in the reverse direction
<i>activity_count</i>	Total number of activities recorded for a given trail segment at a given time in the forward direction
<i>rev_activity_count</i>	Total number of activities recorded for a given trail segment at a given time in the reverse direction
<i>total_activity_count</i>	Total number of activities recorded for a given trail segment at a given time in either direction
<i>activity_time</i>	Total time the activity took while traversing a given trail segment at a given time in the forward direction, in seconds
<i>rev_activity_time</i>	Total time the activity took while traversing a given trail segment at a given time in the reverse direction, in seconds
<i>commute_count</i>	Binary (0 vs. 1) flag to indicate whether or not the activity was recorded as part of a commute

$$v_{\text{Lorentz}} = c \left(\frac{1}{\pi b \left[1 + \left(\frac{\theta - a}{b} \right)^2 \right]} \right) + d + e\theta \quad (14)$$

To best fit the travel rate percentiles for each function Eqs. (12)–(14), the optimal values for coefficients a , b , c , d , and e had to be determined. To do this, the data were first divided in half, at random, into a training dataset ($N_{\text{train}} = 525260$) and a validation dataset ($N_{\text{valid}} = 525261$). For both the training dataset and the validation dataset, the travel rate percentiles (1st, 5th – 95th by 5, and 99th) were

calculated. Using the training dataset, the percentile models were fit using the *nls* (nonlinear least squares) function in R. This function requires the use of starting parameters, which were determined empirically. The resulting predictive models were then compared to the validation dataset percentiles and assessed for fit using R^2 and mean absolute error (MAE).

For comparison to existing travel rate functions, the overall best-fitting slope-travel rate models were applied in a geospatial environment to map and compare estimated travel times. To do this, a single trail within the study area was selected for evaluation. The Lake Blanche Trail, a hiking trail in Big Cottonwood Canyon near Salt Lake

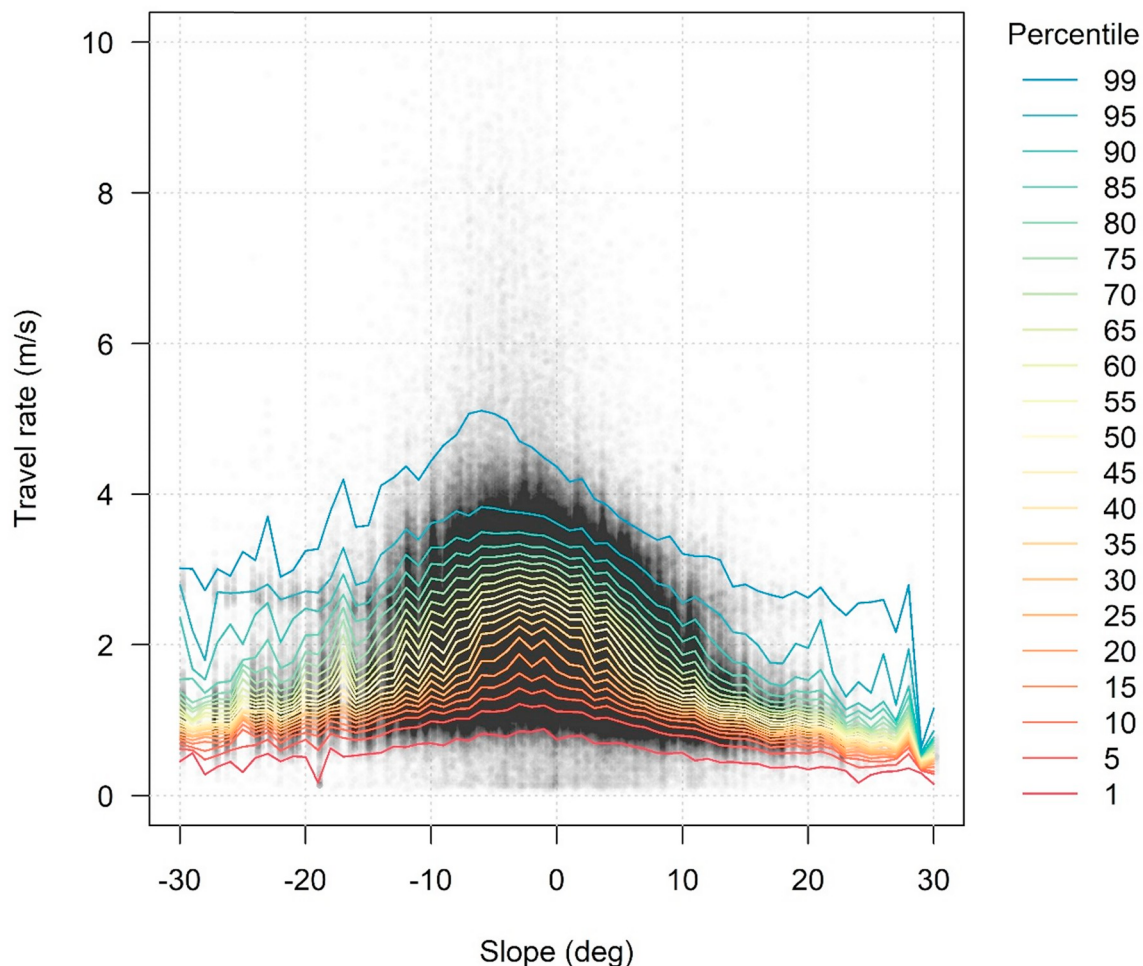


Fig. 4. Raw travel rate vs. slope data with percentiles. Individual travel rate records are represented by black dots with 1% opacity. Thus, dark black areas indicate over 100 overlapping points.

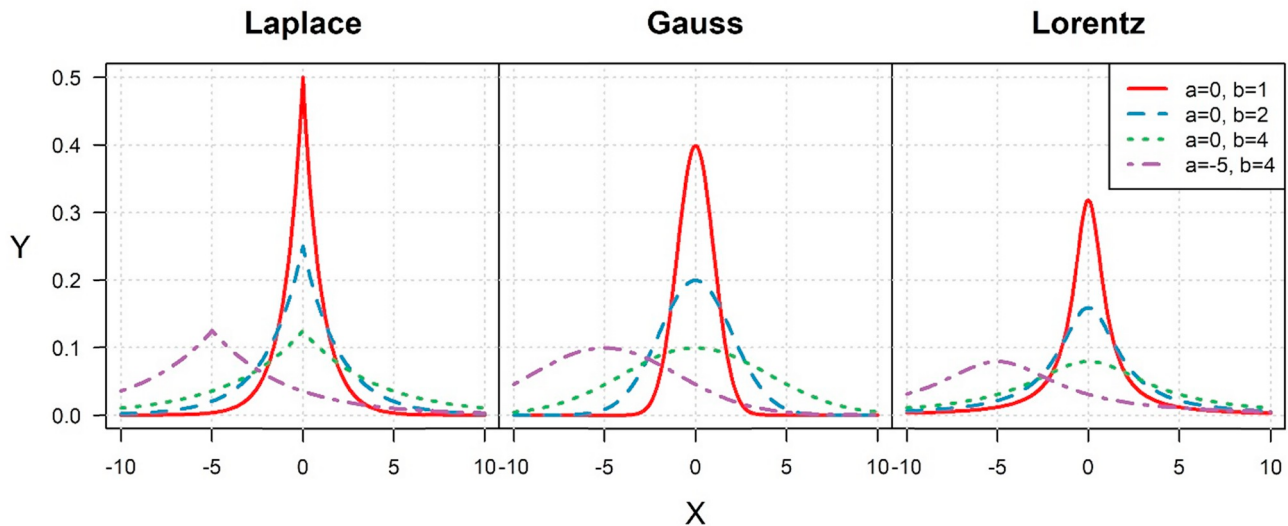


Fig. 5. Laplace, Gauss, and Lorentz function parameter examples.

City, was selected for its popularity, length (~11.7 km, out and back), and diverse terrain (~818 m difference between high and low points). Elevation values were sampled at a 5 m interval along the trail, from which slope (in degrees) was calculated. Slope and horizontal distance were then used to calculate a slope distance. This information was then used to estimate the cumulative travel time for an out and back hike along the entire length of the trail for each of the travel rate percentile models and the existing travel rate functions discussed in the *Background*, with the exception of Davey et al. (1994), since it does not allow for the estimation of downhill travel rates. Lastly, for validation purposes, these modeled results were compared to a sampling of data gathered from AllTrails.com, a popular trail mapping website and mobile application that provides user-submitted GPS tracks of their hikes. We selected the 100 most recent out-and-back hikes of the Lake Blanche Trail (<https://www.alltrails.com/explore/trail/us/utah/lake-blanche-trail>). There was a significant degree of variability in recorded travel distance, which could be caused by a variety of factors, including differences in recorded start- and end-points, GPS error, and venturing off-trail, so to facilitate a more direct comparison, average trail travel rates were calculated for each GPS track and used to estimate total travel time based on the trail length used for modeling.

4. Results

The best-fitting parameters for the three slope-travel rate functions can be seen in Fig. 6. The terms for each model and percentile can be found in the supplemental material. Broadly, the Gauss and Lorentz models tended to be optimized using relatively similar terms, as evident by the similar trends across the range of percentiles. Although there were some similarities in the overall trends, the optimal terms for the Laplace models differed significantly from those of Gauss and Lorentz in magnitude, particularly for the b , c , and d terms. The a term, which dictates the slope of highest travel rate, ranges between -1.41° and -4.00° , a range which is inclusive of the equivalent value for THF (-2.86°), Campbell et al. (2017) (-2.3°), and Irmischer and Clarke (2018) (-2.86°). In each of the three functions, the slope of highest travel rate tends towards steeper downhill slopes from the 5th to the 85th percentile, suggesting that when moving more quickly (e.g. running), one moves fastest at slightly steeper downhill slopes.

In all three cases, the b term, which dictates the width of the curves, begins with a relatively high value, then decreases to a minimum at around the 30th percentile and increases with increasing percentiles. This suggests that, when moving slowly (e.g. 5th) or quickly (e.g. 95th), relatively small increases in slope have a minimal impact on travel

rates. The c term follows a similar pattern to b , generally increasing in all three functions from low to high travel rate percentiles, meaning faster runners have the widest range of travel rates and thus are most affected by slope, whereas the slower hikers have the narrowest range of travel rates and thus are least affected by slope. The d term does not show a very clear trend among the functions and percentiles, as its influence appears to be outweighed by that of c . Lastly, the e term, which dictates the asymmetry of the functions, all have negative values. This suggests that, even after accounting for the fact that people tend to travel fastest on slightly downhill slopes (shifting the curves to the left), there is still an asymmetry to the curve, with uphill travel rates tending to be slightly slower than downhill – a relationship that has not been previously noted or accounted for in predictive models. This asymmetry tends to increase slightly from the 5th to the 85th percentiles, and then sharply from the 85th to the 99th, meaning that the fastest individuals move relatively quickly downhill as compared to uphill travel, whereas the slowest individuals experience less of a difference therebetween.

When modeled using the optimal parameters from Fig. 6, the functions take the forms seen in Fig. 7. As discussed earlier, the Laplace function tends to come to a sharp peak, whereas the Gauss and Lorentz functions tend to produce more rounded maximum travel rates when compared to slope. Thus, Laplace-based slope-travel models tend to produce higher travel rates at the fastest slope as compared to Gauss- and Lorentz-based models. The notable difference in form between the Gauss and Lorentz functions is that Gauss tends to flatten out more in the tails than Lorentz. Accordingly, Gauss-based slope-travel rate models will tend to produce higher travel rates on very steep slopes than Lorentz-based models.

Despite their differences in form, the three functions all produce similar results in terms of model fit (R^2) and prediction error (MAE) (Fig. 8). In general, all three functions perform best when modeling moderately slow to moderately fast travel rates (5th – 85th percentiles) but experience a significant decrease in accuracy at very slow (1st) and fast travel rates (90th – 99th percentiles). The lower accuracy at the 1st percentile is likely due to the influence of immobility on GPS-tracked travel rates, causing significant travel rate variability in the low-end travel rate data. In Fig. 4, particularly on lower slopes, there are significant number of travel rate records at or near zero m/s, which suggests people were likely taking a break in activity, but not stopping their GPS data collection. The lower accuracy at the highest percentiles is likely due to a combination of two factors: (1) the vertical spread in the raw data, particularly on lower slopes exemplified in Fig. 4, likely due to errors in data collection or aggregation; and (2) the fact that running and sprinting travel rates tend to be highly variable between

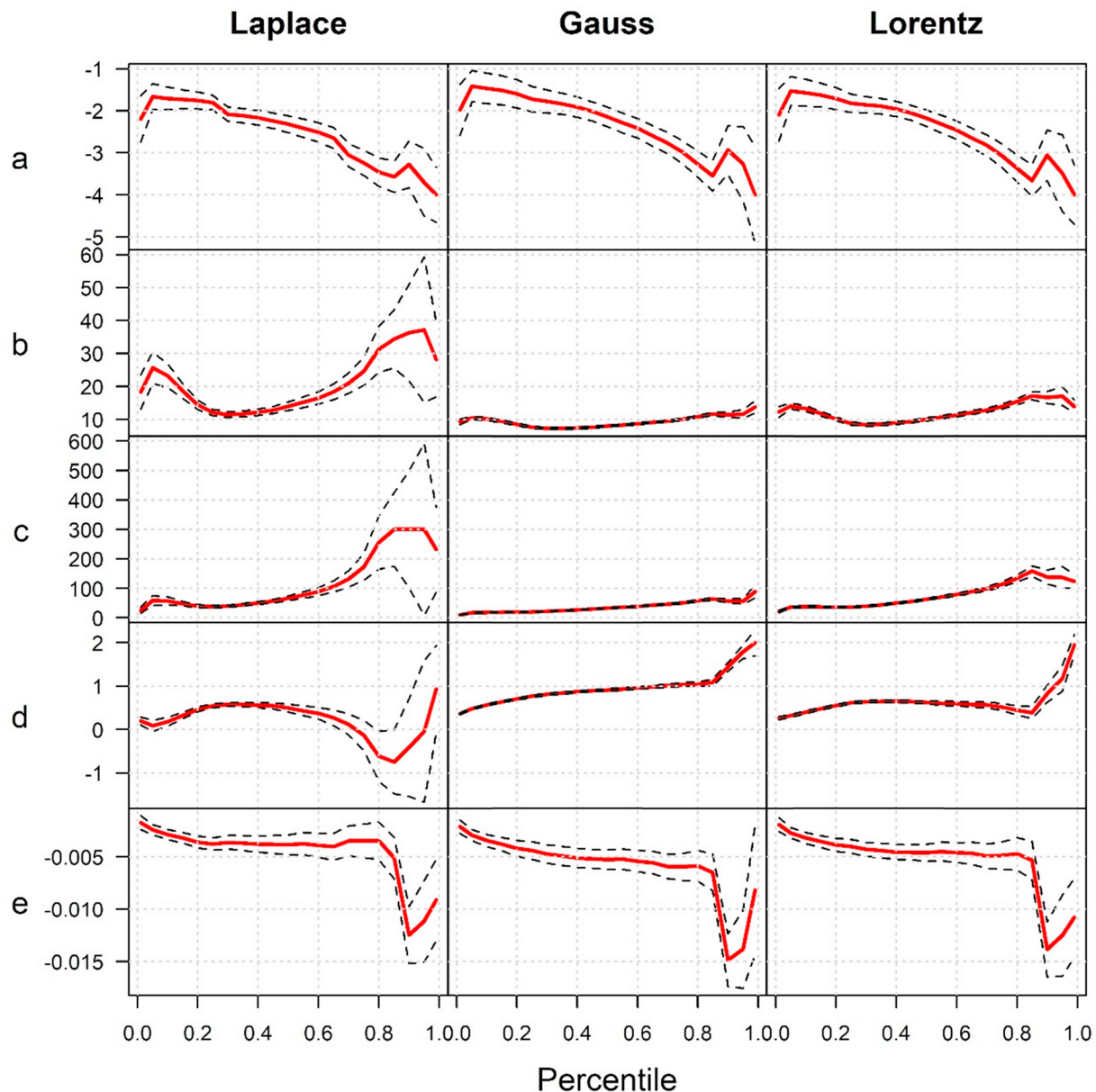


Fig. 6. Optimal terms for travel rate percentile models (1st to 99th) for the Laplace, Gauss, and Lorentz functions, where the solid red line indicates the model term and the dotted black lines indicate the terms' standard errors. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

individuals, especially when compared to walking travel rates. There are some important differences between the models as well. Laplace tends to outperform Gauss and Lorentz at a relatively narrow range of low travel rates (10th – 25th percentiles), whereas Lorentz tends to outperform Laplace and Gauss at a relatively wider range of moderate-fast travel rates (30th – 95th percentiles). Averaged out over the entire range of percentiles tested, Lorentz produced the best results ($R^2_{\text{ave}} = 0.958$; $\text{MAE}_{\text{ave}} = 0.078$ m/s), followed by Laplace ($R^2_{\text{ave}} = 0.953$; $\text{MAE}_{\text{ave}} = 0.088$ m/s), and Gauss ($R^2_{\text{ave}} = 0.949$; $\text{MAE}_{\text{ave}} = 0.090$ m/s).

Notably, the travel rate prediction errors that result from the three functions differ throughout the range of percentiles and the range of slopes (Fig. 9). As mentioned earlier, Laplace tends to outperform Gauss and Lorentz at lower travel rate percentiles. This can also be seen in Fig. 9, where Laplace 10th – 30th percentile model residuals are fairly evenly distributed throughout the range of slopes. However, at higher percentiles, Laplace tends to increasingly and significantly overestimate

travel rates on the fastest (slightly downhill) slopes, as evidenced by the strong black peaks between -5° and 0° . This is due to the peaked nature of the Laplace function. Also at higher percentiles, the sharp decreases in travel rate that occur on either side of the Laplace peak tend to result in underestimation of travel rates, on moderately steep downhill slopes ($\sim -15^\circ$ to -5°) and especially slightly uphill slopes ($\sim 0^\circ$ – 10°). Lastly, at higher percentiles, Laplace tends to overestimate on fairly steep slopes ($\sim -25^\circ$ to -15° ; 10° – 20°), and underestimate on very steep slopes ($\sim < -25^\circ$; $> 20^\circ$).

The Gauss-based models possess very similar patterns of error along the slope axis throughout all of the percentiles tested. They are characterized by the following trends: (1) underestimation at slightly downhill slopes ($\sim -5^\circ$ – 0°), due to the rounded peak of the bell curve; (2) overestimation on either side of the rounded peak ($\sim -15^\circ$ to -5° ; 0° – 10°) due to the width of the rounded peak; (3) underestimation on moderately steep slopes ($\sim -25^\circ$ – -15° ; 10° – 20°) due to the sharp decrease from the rounded peak; and (4) overestimation on very steep

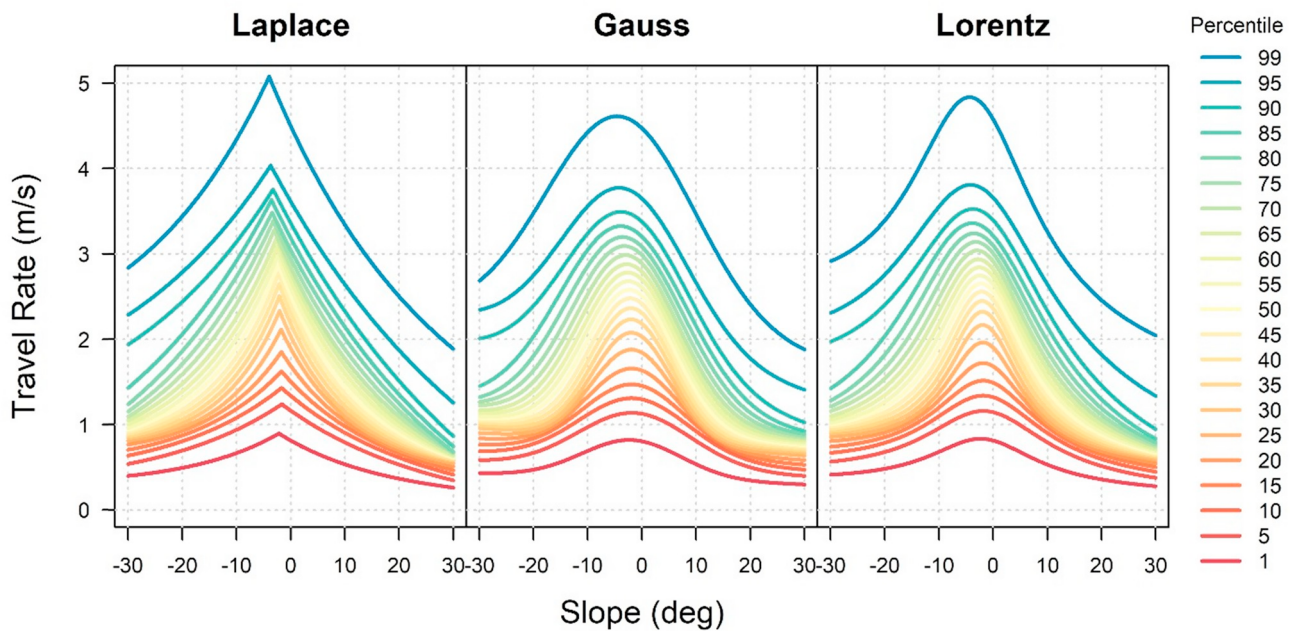


Fig. 7. Best-fitting models predicting travel rate percentiles (1st to 99th) for the Laplace, Gauss, and Lorentz functions based on slope.

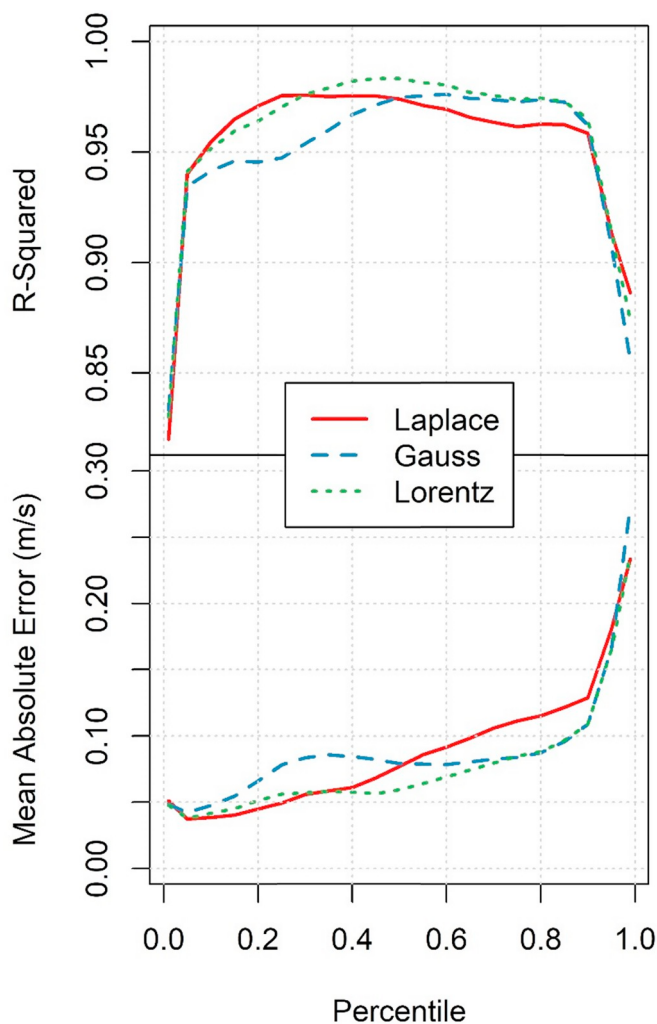


Fig. 8. Accuracy assessment results for the best-fitting models predicting travel rate percentiles (1st to 99th) for the Laplace, Gauss, and Lorentz functions based on slope, including model fit (R^2) and mean absolute error.

slopes ($\sim < -25^\circ$; $> 20^\circ$) due to the flattening of the tails.

The Lorentz-based model residuals are similar in form to those of the Gauss-based models, but the magnitude of their over- and under-estimations are generally lower. In fact, it appears that the Lorentz models tend to balance out the model residuals throughout the range of slopes the best out of the three.

Although the Strava data contained no mention of the specific mode of travel (e.g. walking vs. jogging vs. running), and no physiological information to suggest level of energy exertion (e.g. heartrate), a comparison of our modeled results to existing travel rate functions revealed notable similarity between the results of other, walking/hiking-based studies and the lower travel rate percentiles from our study (Fig. 10). None of the existing functions matches any of the individual Lorentz-based modeled percentiles exactly, but that of [Irmischer and Clarke \(2018\)](#) comes the closest to the 5th percentile. This is likely due to their large sample size ($N = 200$) and the use of a Gaussian function, which is similar in form to Lorentz. With the exception of the anomalous [Langmuir \(1984\)](#) corrections, [Campbell et al. \(2017\)](#)'s travel rates are higher than any of the other functions for the entire slope range, even though they were based on a walking experiment. This is likely due to the very short distances traversed by study subjects (100 m), the associated lack of a fatigue effect, and the quadratic function used to model slope effects. The functions of [Tobler \(1993, p. 24\)](#), [Naismith \(1892\)](#), [Rees \(2004\)](#) all tend to underestimate walking travel rates on steeper slopes.

Fig. 11 illustrates the results of the travel time simulation on the Lake Blanche Trail. According to these simulations, moving at the slowest modeled rate (1st percentile) results in a total travel time of approximately 5.5 h, which is over 6 times longer than that of the fastest rate (99th percentile), at approximately 0.9 h. With the exception of [Campbell et al. \(2017\)](#), most of the existing travel rate functions resulted in travel times similar to that of the Lorentz-based 5th percentile model, with [Tobler \(1993, p. 24\)](#), [Naismith \(1892\)](#)-[Langmuir \(1984\)](#), and [Rees \(2004\)](#) resulting in slightly longer travel times than the 5th percentile, and [Irmischer and Clarke \(2018\)](#) resulting in a slightly shorter travel time. [Campbell et al. \(2017\)](#) is a clear outlier, with a resultant travel time that closely approximates the Lorentz-based 45th percentile model. This is probably due to the fact that the [Campbell et al. \(2017\)](#) experiments were conducted on short, 100 m transects. Thus, the study subjects were prone to faster travel than they

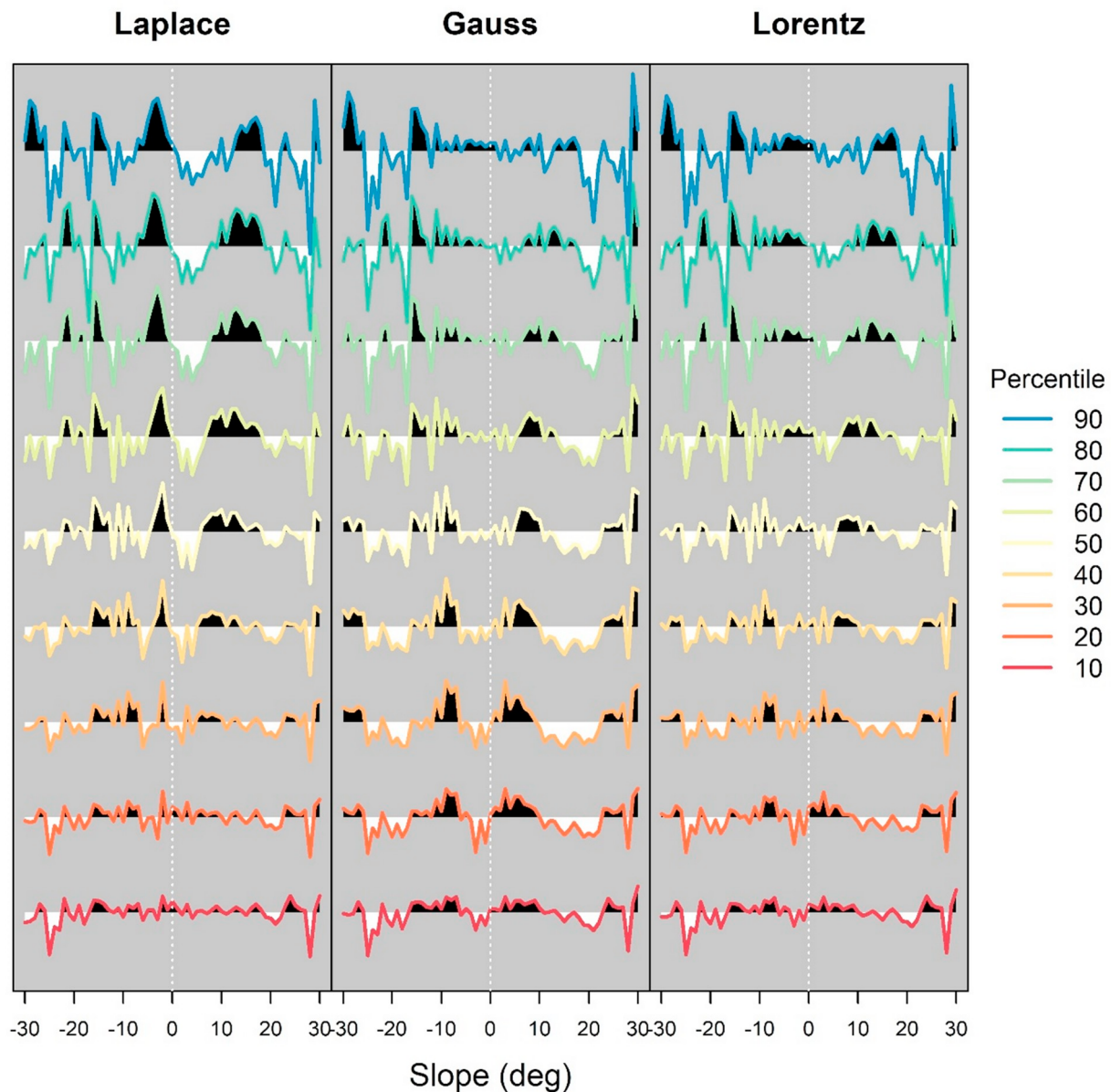


Fig. 9. Model residuals for a selection of the best-fitting models predicting travel rate percentiles (10th – 90th, by 10) for the Laplace, Gauss, and Lorentz functions based on slope. Black areas represent model overestimation; white areas represent model underestimation.

would on a longer hike.

Existing travel rate functions do not quantify the significant variability in travel rates within a population. Fig. 12 demonstrates variability in walking/hiking travel rates, representing a distribution of total travel times of real hiking data gleaned from the Lake Blanche Trail GPS tracks ($n = 100$) found on AllTrails.com. Although all of the tracks were self-reported to be classified as “hiking” (rather than other AllTrails.com categories such as “backpacking” or “trail running”), the resulting travel times ranged widely from 1.9 h to 5.6 h. These extremes corresponded most closely with the simulated travel times of the 55th and 1st percentiles from the Lorentz-based models, respectively (Fig. 12). The mean hiking travel time from the validation data was about 3.7 h, which was closely approximated by both the Lorentz-based 5th percentile model and Irmischer and Clarke (2018), with the former predicting slightly slower travel, and the latter predicting slightly faster. According to these data, the Lorentz-based 1st percentile could be used as a representative for the extremely slow end of the travel rate spectrum, the 5th percentile could be used as a representative for average hiking rates, and, depending on level of energy exertion,

anywhere between the 10th and 55th percentiles could be used to represent increasingly-fast hiking.

5. Discussion

The slope-travel rate models we have presented in this research are based on the most comprehensive database of travel rate records known to the authors in the scientific literature at the time of writing. As such, we have provided perhaps the most generalizable and broadly-applicable set of functions for estimating travel rates and times along trails to date. We have presented these models based on three different mathematical formulations, which, although statistically-similar, can produce notably different travel rate and time estimates, particularly over longer distances. To predict walking travel rate as a function of slope we suggest applying the Laplace and/or Lorentz models shown in Eqns. (12) and (14), respectively, along with the optimal model terms found in the supplemental material. If a single, flexible travel rate function is desired, then we suggest applying the Lorentz function Eqn. (14), given its superior performance at the 5th percentile, which most closely

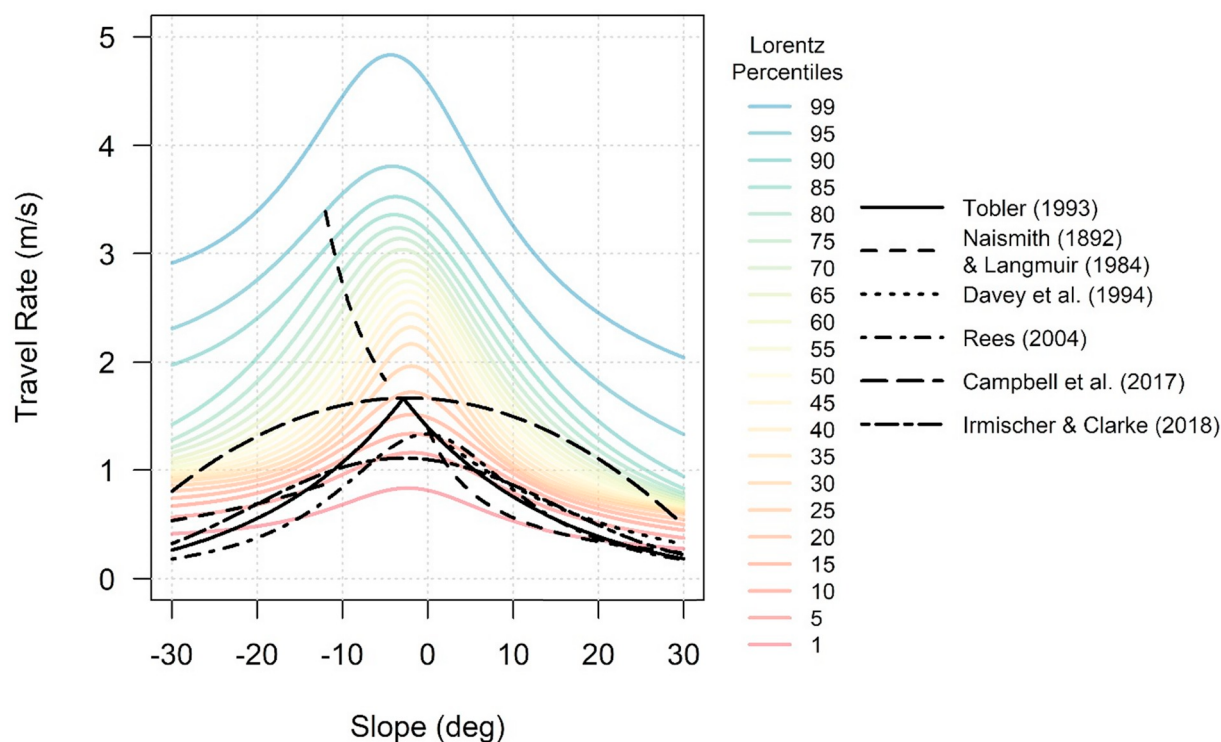


Fig. 10. Comparison of the Lorentz-based travel rate percentile models to the existing travel rate functions depicted in Fig. 1.

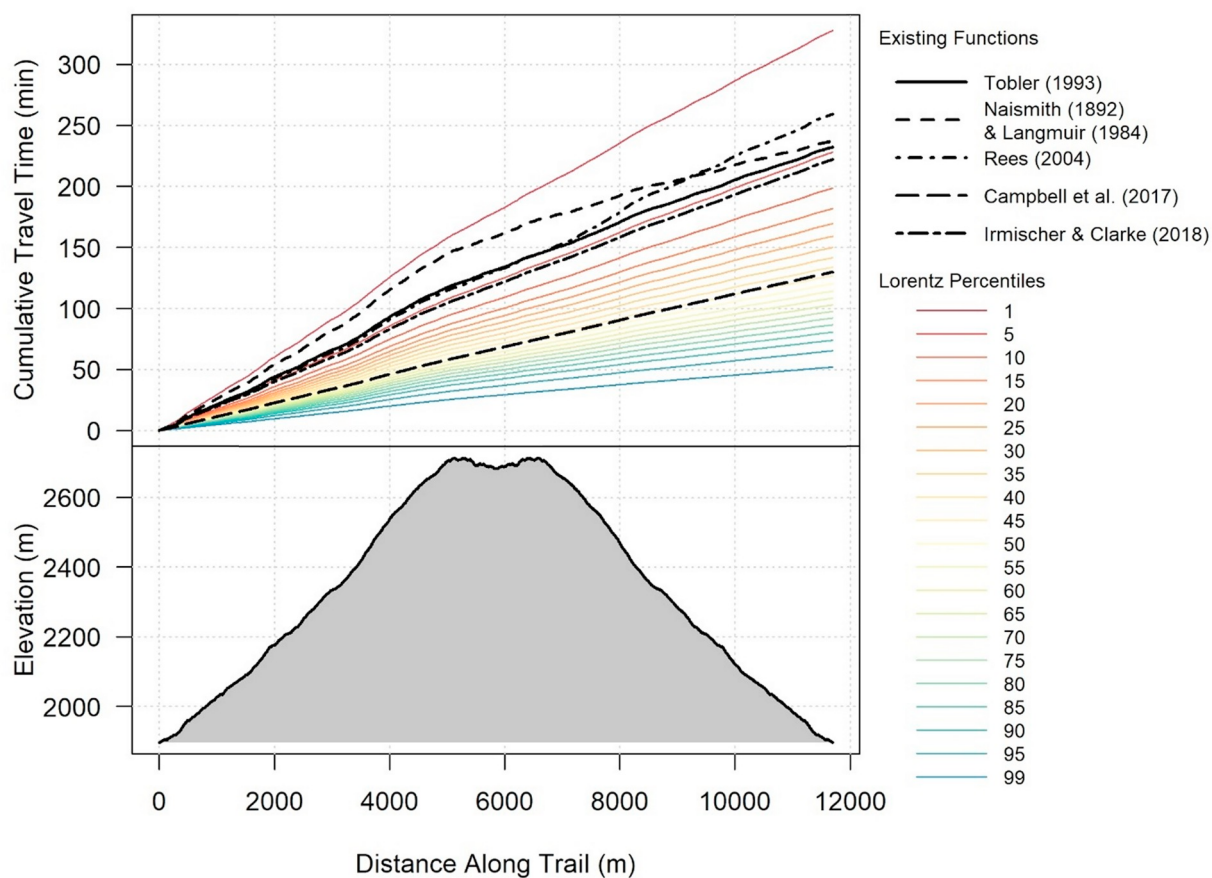


Fig. 11. Comparison of simulated travel times for hiking the Lake Blanche Trail out and back with elevation profile.

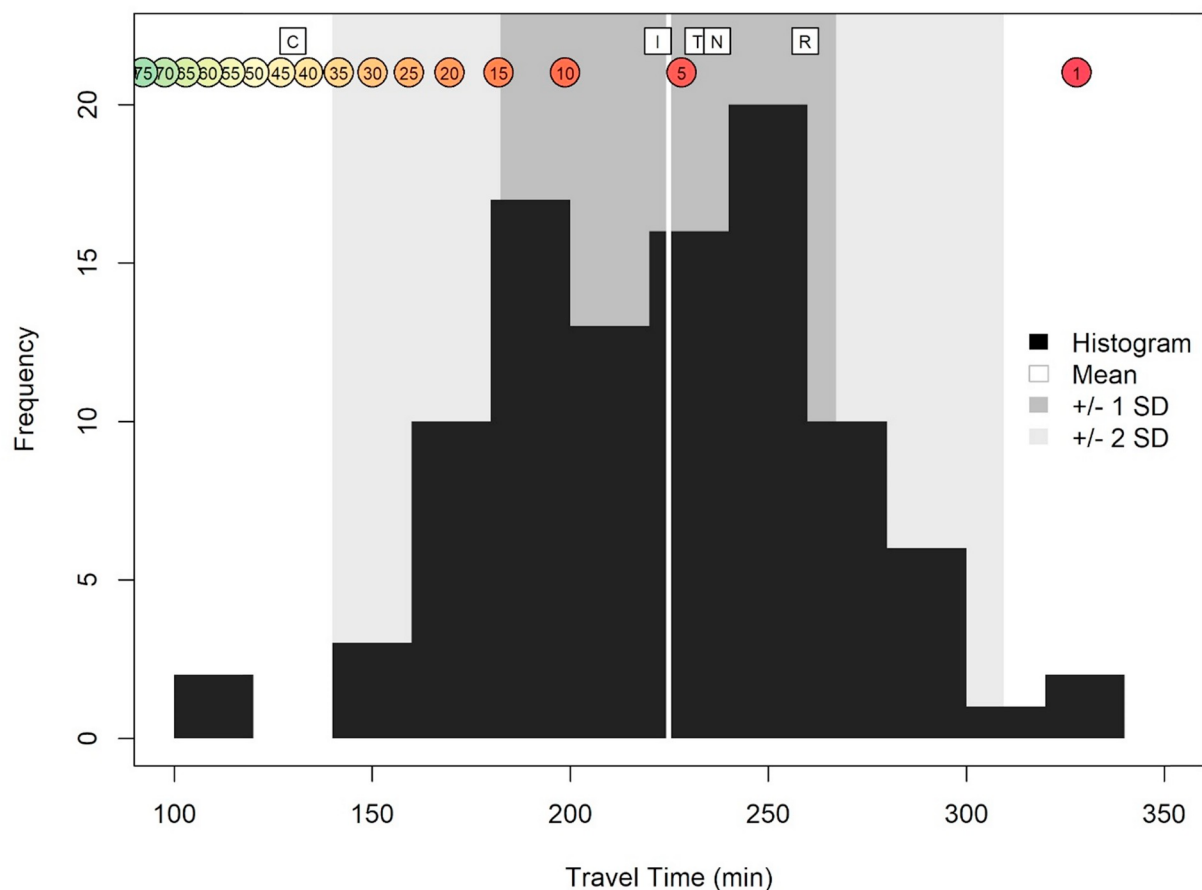


Fig. 12. Histogram of the Lake Blanche Trail hike travel times from selection of GPS tracks ($n = 100$) obtained from [AllTrails.com](https://www.alltrails.com) as compared to simulated travel times. Circles with numbers represent Lorentz-based travel rate percentile results. Squares with letters represent existing travel rate functions (C = [Campbell et al. \(2017\)](#); I = [Irmischer and Clarke \(2018\)](#); T = [Tobler \(1993, p. 24\)](#); N = [Naismith \(1892\) with Langmuir \(1984\)](#); R = [Rees \(2004\)](#)).

matches the existing travel rate functions ([Fig. 11](#)) and the average hiking rate from our validation dataset ([Fig. 12](#)), as well as its superior performance at higher percentiles (30th – 95th). At moderately low travel rates (10th – 25th percentile), Laplace ([Eqn. \(12\)](#)) appears to produce the most reliable predictions.

One of the main strengths of this research – the use of a massive database of crowdsourced GPS-tracked travel rates – is also a source of several potential complications. For example, the anonymity of the data we received from Strava does not allow for a nuanced demographic analysis. Accordingly, we were unable to generate a separate function for males vs. females, younger users vs. older users, high vs. low body mass index, etc. Individuals may have a difficult time understanding where they fit relative to an anonymous body of Strava users, thus limiting their ability to confidently place themselves into one percentile or another. However, instead of trying to fit precisely into a percentile, these models can be used to present a range of travel rates and times. For example, if a wildland firefighter were using a mobile application geared towards estimating time to safety ([Campbell et al., 2017](#)), perhaps he or she would want to err on the side of caution, and assume a 1st–10th percentile travel rate. Similarly, if a search and rescue team was trying to estimate the time it would take to reach an injured hiker with a time-sensitive, life-threatening injury ([Ciesa, Grigolato, & Cavalli, 2014](#)), they would be able to estimate a range of possible arrival times (e.g. best case scenario at 99th percentile; worst-case scenario at 1st percentile) and determine the chances of survivability. Alternatively, if an elite trail runner were trying to predict their race results, perhaps he or she would want to use a high percentile model better calibrated to their level of athletic ability. In addition to these examples, our hope is that in any of the existing application areas in

which terrain-based least cost path modeling is employed (e.g. archaeology, urban evacuation, social network analysis), the improved slope-travel rate effects that we have modeled will improve the accuracy of simulated human movement.

Another shortcoming of the anonymity of the database is the fact that we have no identifying information to determine how far someone's entire hike or run was; we only have the segment of trail and the time traversed on a given segment. We know that the median travel distance and time were about 6.5 km and 49 min, respectively, among the entire study population, but we do not have these data on a per-activity basis. This is problematic because distance traveled can affect travel rates. Fatigue will reduce travel rate over time. Also, someone running a trail marathon would run more slowly than someone sprinting up a short section of trail, to maintain a consistent level of energy expenditure.

Although this research is based on a very large database of GPS tracks drawn from a wide range of activities performed by an age- and gender-diverse population ([Table 3](#)), it is important to consider the degree to which data from a fitness tracking application such as Strava are representative of “normal” pedestrian activities. As [Barratt \(2017\)](#) points out, Strava, and other GPS-based fitness tracking, social media-driven mobile applications can introduce a level of competition between users, which can act to increase the level of intensity with which they pursue a given activity. In effect, this can result in an upward bias of travel rate data gleaned from these applications. Thus, particularly for the higher-percentile travel rates, it is certainly possible that our predictive models are representative of a higher-than-average level of exertion. In the absence of associated physiological data, such as heart rate and/or oxygen consumption, there is no way to normalize travel

rates for exertion level.

The data we received were noisy, as is evident in Fig. 4 by looking at the vertical spread in travel rate records, particularly on low slopes. We attempted to remove outliers by cutting travel rates off at 10 m/s, but it is still very likely that some records are erroneous. There could be many causes for these errors, including GPS error, user error, and application error, but unfortunately there was no way to determine the specific causes and remove them accordingly. However, given the size of the database and the vast number of non-erroneous records contained therein, we believe that these outliers bore little effect on the modeling results.

There are, of course, many variables that affect travel rates other than slope, some of which are intrinsic (e.g. fitness level) and impossible to study using an anonymous database, but others of which are extrinsic and potentially able to be studied using our dataset. For example, variables such as elevation, time of day, time of year, temperature, humidity, precipitation, ground surface stability and roughness, vegetation density, and others could potentially be explored in the future using similar data.

6. Conclusions

In this paper we have presented a statistically-robust and flexible set of functions for predicting pedestrian travel rates as a function of terrain slope based on a vast, crowdsourced database of GPS tracks compiled and provided by the popular fitness application, Strava. We aim for these functions to serve as a reliable and accurate basis of travel rate and time estimation moving forward, applicable in any of the abundant disciplines in which terrain effects are of interest, be it the estimation of the time it will take for a wildland fire crew to access a safety zone, a search and rescue crew to reach an accident victim, or a trail runner to complete a race. Although widely-cited, the shortcomings of the scientific approaches that led to THF and Naismith's Rule, in particular, may render them less effective than the broadly-applicable set of functions we have presented. Specifically, if a single travel rate function is desired across the gamut of travel modes, we recommend applying the Lorentz-based models, with the 5th percentile being most appropriate for hiking travel rates. Higher-percentile models should be used for applications where a higher level of physical exertion (e.g. jogging, running, sprinting) is anticipated, but specific percentiles must be selected based on a careful calibration to one's own travel rate. To build upon our work, future research should aim to quantify important variables such as load carriage, travel distance fatigue, and link the travel rate percentiles back to quantifiable measures of energy expenditure and mode of travel.

Declarations of interest

None.

Acknowledgments

Funding for this research was provided by the USDA Forest Service National Fire Plan through the Office of Research, and the National Wildfire Coordinating Group Fire Behaviour Subcommittee, Cooperative Agreements 15CR11221637105, 18JV11221637153, and 18JV11221637154. This work was also supported by National Science Foundation grant number DEB-1714972.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.apgeog.2019.03.008>.

References

- Alegana, V. A., Wright, J. A., Pentrina, U., Noor, A. M., Snow, R. W., & Atkinson, P. M. (2012). Spatial modelling of healthcare utilisation for treatment of fever in Namibia. *International Journal of Health Geographics*, 11, 6. <https://doi.org/10.1186/1476-072X-11-6>.
- Alexander, M. E., Baxter, G. J., & Dakin, G. R. (2005). Travel rates of Alberta wildland firefighters using escape routes. In B. W. Butler, & M. E. Alexander (Eds.). *Human factors - 10 Years later* Missoula, MT: International Association of Wildland Fire. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.535.9337&rep=rep1&type=pdf>.
- Barratt, P. (2017). Healthy competition: A qualitative study investigating persuasive technologies and the gamification of cycling. *Health & Place*, 46, 328–336. <https://doi.org/10.1016/j.healthplace.2016.09.009>.
- Butler, B. W., Cohen, J. D., Putnam, T., Bartlette, R. A., & Bradshaw, L. S. (2000). A method for evaluating the effectiveness of firefighter escape routes. *4th international wildland fire safety summit* (pp. 10–12).
- Campbell, M. J., Dennison, P. E., & Butler, B. W. (2017). A LiDAR-based analysis of the effects of slope, vegetation density, and ground surface roughness on travel rates for wildland firefighter escape route mapping. *International Journal of Wildland Fire*, 26(10), 884–895. <https://doi.org/10.1071/WF17031>.
- Carver, S., Comber, A., McMorran, R., & Nutter, S. (2012). A GIS model for mapping spatial patterns and distribution of wild land in Scotland. *Landscape and Urban Planning*, 104(3), 395–409. <https://doi.org/10.1016/j.landurbplan.2011.11.016>.
- Chiou, C.-R., Tsai, W.-L., & Leung, Y.-F. (2010). A GIS-dynamic segmentation approach to planning travel routes on forest trail networks in Central Taiwan. *Landscape and Urban Planning*, 97(4), 221–228. <https://doi.org/10.1016/j.landurbplan.2010.06.004>.
- Ciesia, M., Grigolato, S., & Cavalli, R. (2014). Analysis on vehicle and walking speeds of search and rescue ground crews in mountainous areas. *J. Outdoor Recreat. Tour.* 5(6), 48–57. <https://doi.org/10.1016/j.jort.2014.03.004>.
- Contreras, D. A. (2011). How far to conchucos? A GIS approach to assessing the implications of exotic materials at chavín de Huántar. *World Archaeology*, 43(3), 380–397. <https://doi.org/10.1080/00438243.2011.605841>.
- Davey, R. C., Hayes, M., & Norman, J. M. (1994). Running uphill: An experimental result and its applications. *Journal of the Operational Research Society*, 45(1), 25–29. <https://doi.org/10.2307/2583947>.
- De Silva, C. S., Warusavitharana, E. J., & Ratnayake, R. (2017). An examination of the temporal effects of environmental cues on pedestrians' feelings of safety. *Computers, Environment and Urban Systems*, 64, 266–274. <https://doi.org/10.1016/j.compenurbysys.2017.03.006>.
- Doherty, P. J., Guo, Q., Doka, J., & Ferguson, D. (2014). An analysis of probability of area techniques for missing persons in Yosemite National Park. *Applied Geography*, 47, 99–110. <https://doi.org/10.1016/j.apgeog.2013.11.001>.
- Fryer, G. K., Dennison, P. E., & Cova, T. J. (2013). Wildland firefighter entrapment avoidance: Modelling evacuation triggers. *International Journal of Wildland Fire*, 22(7), 883–893. <https://doi.org/10.1071/WF12160>.
- Gravel-Miguel, C. (2016). Using Species Distribution Modeling to contextualize Lower Magdalenian social networks visible through portable art stylistic similarities in the Cantabrian region (Spain). *Quaternary International*, 412, 112–123. <https://doi.org/10.1016/j.quaint.2015.08.029>.
- Heesch, K. C., & Langdon, M. (2017). The usefulness of GPS bicycle tracking data for evaluating the impact of infrastructure change on cycling behaviour. *Health Promotion Journal of Australia*, 27(3), 222–229. <https://doi.org/10.1071/HE16032>.
- Henry, D. O., Belmaker, M., & Bergin, S. M. (2017). The effect of terrain on Neanderthal ecology in the Levant. *Quaternary International*, 435, 94–105. <https://doi.org/10.1016/j.quaint.2015.10.023>.
- Herzog, I. (2010). Theory and practice of cost functions. In F. Contreras, M. Farjas, & F. Melero (Eds.). *Fusion of cultures*. Granada, Spain: Archaeopress.
- Imhof, E. (1950). *Gelände und Karte*. Erlenbach-Zürich: E. Rentsch.
- Irmischer, I. J., & Clarke, K. C. (2018). Measuring and modeling the speed of human navigation. *Cartography and Geographic Information Science*, 45(2), 177–186. <https://doi.org/10.1080/15230406.2017.1292150>.
- Jennings, J., & Craig, N. (2001). Politywide analysis and imperial political economy: The relationship between valley political complexity and administrative centers in the wari empire of the central andes. *Journal of Anthropological Archaeology*, 20(4), 479–502. <https://doi.org/10.1006/jaar.2001.0385>.
- Jensen, D. E. (2003). Geoglyphs and GIS: Modeling transhumance in northern Chile. In M. Doerr, & S. Apostolis (Eds.). *The digital heritage of archaeology* (pp. 6). Heraklion, Crete.
- Kantner, J. (2004). Geographical approaches for reconstructing past human behavior from prehistoric roadways. In M. F. Goodchild, & D. G. Janelle (Eds.). *Spatially integrated social science* (pp. 323–344). New York, NY: Oxford University Press.
- Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly landsat time series: 1. LandTrendr — temporal segmentation algorithms. *Remote Sensing of Environment*, 114(12), 2897–2910. <https://doi.org/10.1016/j.rse.2010.07.008>.
- Kunitz, J. K., Lagree, J. D., & Weinig, D. L. (2017). A GIS examination of the chacoan Great North road. *KIVA*, 83(1), 86–113. <https://doi.org/10.1080/00231940.2016.1199936>.
- Langmuir, E. (1984). *Mountain craft and leadership: A handbook for mountaineers and hill walking leaders in the British Isles*. Scottish Sports Council Edinburgh.
- Márquez-Pérez, J., Vallejo-Villalta, I., & Álvarez-Francisco, J. I. (2017). Estimated travel time for walking trails in natural areas. *Geogr. Tidsskr.-Dan. J. Geogr.* 117(1), 53–62. <https://doi.org/10.1080/00167223.2017.1316212>.

- McCoy, M. D., Mills, P. R., Lundblad, S., Rieth, T., Kahn, J. G., & Gard, R. (2011). A cost surface model of volcanic glass quarrying and exchange in Hawai'i. *Journal of Archaeological Science*, 38(10), 2547–2560. <https://doi.org/10.1016/j.jas.2011.04.017>.
- Minetti, A. E., Moia, C., Roi, G. S., Susta, D., & Ferretti, G. (2002). Energy cost of walking and running at extreme uphill and downhill slopes. *Journal of Applied Physiology*, 93(3), 1039–1046. <https://doi.org/10.1152/japplphysiol.01177.2001>.
- Mink, P. B., Ripy, J., Bailey, K., & Grossardt, T. H. (2009). Predictive archaeological modeling using GIS-based fuzzy set estimation: A case study in woodford county, Kentucky. *Proceedings of the ESRI user conference* (pp. 18). (San Diego, CA).
- Musakwa, W., & Selala, K. M. (2016). Mapping cycling patterns and trends using Strava Metro data in the city of Johannesburg, South Africa. *Data Brief*, 9, 898–905. <https://doi.org/10.1016/j.dib.2016.11.002>.
- Naismith, W. (1892). Cruach adran, stobinian, and ben more. *Scott. Mountaineering Club J.* 2(3), 136.
- Ranacher, P., Brunauer, R., Trutschnig, W., Spek, S. V. der, & Reich, S. (2016). Why GPS makes distances bigger than they are. *International Journal of Geographical Information Science*, 30(2), 316–333. <https://doi.org/10.1080/13658816.2015.1086924>.
- Rees, W. G. (2004). Least-cost paths in mountainous terrain. *Computers & Geosciences*, 30(3), 203–209. <https://doi.org/10.1016/j.cageo.2003.11.001>.
- Richards-Rissetto, H., & Landau, K. (2014). Movement as a means of social (re)production: Using GIS to measure social integration across urban landscapes. *Journal of Archaeological Science*, 41, 365–375. <https://doi.org/10.1016/j.jas.2013.08.006>.
- Ripy, J., Grossardt, T., Shouse, M., Mink, P., Bailey, K., & Shields, C. (2014). Expert systems archeological predictive model, expert systems archeological predictive model. *Transportation Research Record*, 2403(1), 37–44. <https://doi.org/10.3141/2403-05>.
- Rogers, S. R., Collet, C., & Lugon, R. (2014a). Least cost path analysis for predicting glacial archaeological site potential in central Europe. In A. Travaglia (Ed.), *Across space and time* (pp. 261). Perth: Amsterdam University Press.
- Rogers, S. R., Fischer, M., & Huss, M. (2014b). Combining glaciological and archaeological methods for gauging glacial archaeological potential. *Journal of Archaeological Science*, 52, 410–420. <https://doi.org/10.1016/j.jas.2014.09.010>.
- Strava (2018). 2017 in stats. Retrieved July 9, 2018, from <https://blog.strava.com/2017-in-stats/>.
- Sun, Y., & Mobasheri, A. (2017). Utilizing crowdsourced data for studies of cycling and air pollution exposure: A case study using Strava data. *International Journal of Environmental Research and Public Health*, 14(3), 274. <https://doi.org/10.3390/ijerph14030274>.
- Taliaferro, M. S., Schriever, B. A., & Shackley, M. S. (2010). Obsidian procurement, least cost path analysis, and social interaction in the Mimbres area of southwestern New Mexico. *Journal of Archaeological Science*, 37(3), 536–548. <https://doi.org/10.1016/j.jas.2009.10.018>.
- Tobler, W. (1993). *Three presentations on geographical analysis and modeling* (No. 93–1). University of California at Santa Barbara: National Center for Geographic Information and Analysis.
- Tomppo, E., Malimbwi, R., Katila, M., Mäkisara, K., Henttonen, H. M., Chamuya, N., et al. (2014). A sampling design for a large area forest inventory: Case Tanzania. *Canadian Journal of Forest Research*, 44(8), 931–948. <https://doi.org/10.1139/cjfr-2013-0490>.
- United States Census Bureau (2010). *TIGER/Line shapefile 2010 Census Block state-based*. Retrieved from <http://www.census.gov/geo/www/tiger>.
- Utah Geological Survey. (2000). Glad you asked: How was Utah's topography formed? – Utah geological Survey. Retrieved January 28, 2019, from <https://geology.utah.gov/map-pub/survey-notes/glad-you-asked/how-was-utahs-topography-formed/>.
- Verhagen, P., & Jeneson, K. (2012). A Roman puzzle. Trying to find the via Belgica with GIS. *Thinking beyond the tool. Archaeological computing and the interpretive process* (pp. 123–130). Oxford: Archaeopress.
- White, D. A. (2015). The basics of least cost analysis for archaeological applications. *Adv. Archaeol. Pract.* 3(4), 407–414. <https://doi.org/10.7183/2326-3768.3.4.407>.
- Wood, N. J., & Schmidtlein, M. C. (2012). Anisotropic path modeling to assess pedestrian-evacuation potential from Cascadia-related tsunamis in the US Pacific Northwest. *Natural Hazards*, 62(2), 275–300. <https://doi.org/10.1007/s11069-011-9994-2>.