Route choice in mountain navigation, Naismith's rule, and the equivalence of distance and climb

To cite this Article: Route choice in mountain navigation, Naismith's rule, and the equivalence of distance and climb, Journal of Sports Sciences, 25:6, 719 - 726
URL: http://dx.doi.org/10.1080/02640410600874906

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

© Taylor and Francis 2007
Route choice in mountain navigation, Naismith’s rule, and the equivalence of distance and climb

PHILIP SCARF

Centre for OR and Applied Statistics, University of Salford, Salford, UK

(Accepted 28 April 2006)

Abstract
In this paper, I consider decision making about routes in mountain navigation. In particular, I discuss Naismith’s rule, a method of calculating journey times in mountainous terrain, and its use for route choice. The rule is essentially concerned with the equivalence, in terms of time duration, between climb or ascent and distance travelled. Naismith himself described a rule that is purported to be based on trigonometry and simple assumptions about rate of ascent; his rule with regard to hill-walking implies that 1 m of ascent is equivalent to 7.92 m of horizontal travel (1:7.92). The analysis of data on fell running records presented here supports Naismith’s rule and it is recommended that male runners and walkers use a 1:8 equivalence ratio and females a 1:10 ratio. The present findings are contrasted with those based on the analysis of data relating to treadmill running experiments (1:3.3), and with those based on the analysis of times for a mountain road-relay (1:4.4). Analysis of cycling data suggests a similar rule (1:8.2) for cycling on mountainous roads and tracks.

Keywords: Naismith’s rule, regression analysis, extreme values

Introduction
Choosing between alternative routes is a key part of mountain navigation and good decisions about routes is a requisite for success in mountain navigation sports such as mountain marathons, orienteering, adventure racing, and to some extent fell running. In this paper, I focus on route choice in mountain marathons. Such events typically take place over 2 days in mountainous terrain; competitors travel on foot, navigating from point to point and carrying food and equipment necessary to be self-sufficient at an overnight camp. Distances involved can range from 40 to 90 km in total depending on the event and the class. For reasons of safety in the mountains and extreme terrain, most competitors run in pairs, and the fastest team (pair) over the course wins. Broadly speaking, there are three aspects to competing in such events: route choice, fine navigation, and fitness. That is, deciding which way to go, finding the checkpoint on approach, and moving quickly through the terrain encountered on the route. One could argue that a fourth and important aspect is “flow” – performing all tasks calmly and efficiently while maintaining progression. Route choice – deciding which way to go – is a matter of identifying potential routes and choosing the fastest route among them. For example, Figure 1 shows a typical leg on a mountain marathon; there is no restriction on the route to be taken, only that runners must visit checkpoint 2 after checkpoint 1. Thus competitors are faced with the decision about whether: (1) to go directly via the steep slopes of Red Pike; (2) to go the south, with fast running on the road; or (3) to take a northern route along the ridge, staying high. Such route choice has to be made “on the run”, as maps are only made available at the start.

For many years hill walkers and fell runners have been using Naismith’s rule as a basis for estimating journey times. In an article for the Scottish Mountaineering Club, Naismith (1892) describes an expedition he undertook on 2 May 1892 up Cruach Ardran, Stobinian, and Ben More in the Scottish Highlands (“distance, ten miles; total climb, 6,300 feet”). He concludes the article by stating that his journey time of $6\frac{1}{2}$ h, including short rests, “tallies exactly with a simple formula that may be found useful in estimating what time men in fair condition should allow for easy expeditions, namely, an hour for every three miles on the map with an additional hour for every 2000 ft of ascent”. Scarf (1998)
Figure 1. Saunders Lakeland Mountain Marathon, 2001: 1–2, a long-leg on day 1. Map of the Western Lake District with kilometre squares and 15-m contour interval © HARVEY 2006. Grid reference for checkpoint 1 is NY224122. Three routes indicated: —— , road route; —— , middle route; ——— , Pillar route.
discussed this rule and its empirical basis in a short article for a mathematics newsletter, but a complete analysis of the data was not presented. A complete analysis is presented here and the results obtained are contrasted with the findings of other empirical studies that have investigated Naismith’s rule. How this rule is used in its equivalence form for evaluating competing routes is then illustrated. While mountain marathon competitors will be concerned solely with the duration of a route (which route is quickest?), walkers are likely to be concerned more with the aesthetics of routes, and would therefore be more likely to choose the high level ridge (northern route) in Figure 1. However, quantitative decision support for route choice will still be useful for them. A similar rule has been developed for cyclists (Scarf & Grehan, 2005); this rule has implications for navigation in mountain-bike orienteering races.

It has been claimed by Mills (1982) that Naismith’s rule goes back to some ideas of the eighteenth-century Scottish mathematician, Colin MacLaurin. Correspondence between MacLaurin and a fellow member of the Edinburgh Philosophical Society in the 1740s was concerned with the problem of how much a man could raise working a treadmill. Such man-powered treadmills were commonplace. For example, the Harwich treadwheel crane (Harwich Society, 2005) shown in Figure 2 was built in 1667 and remained in use until 1928 in the Naval Yard at Harwich, Essex, UK – the crane was worked by men walking in the interior of each of two wheels 16 ft in diameter. In his response to his correspondent, MacLaurin supposed “that a man can walk on a plane inclined to the horizon in an angle of 30 degrees at a rate of a foot or rather a little more in a second”; a little trigonometry implies therefore “that a man can raise his weight something more than 1800 ft in an hour”. It is unknown whether Naismith knew of this when he devised the rule some 150 years later.

Since fitness varies among individuals, the rule is best used to consider the equivalence, on the basis of time taken, of climb and flat distance; so that 2000 ft of ascent (climb) is equivalent to 3 miles of horizontal travel (distance). Thus 1 unit of climb is equivalent to 7.92 units of distance (1:7.92). I refer to the equivalence factor, 7.92 in this case, as Naismith’s number, and denote it by $x$. More detailed rules have been suggested to account for steepness of the ground and prevailing conditions. Langmuir (1984) states that 10 min should be subtracted for every 300 m of descent on slopes between $5^\circ$ and $12^\circ$, and that 10 min should be added for every 300 m of descent on slopes greater than $12^\circ$. A similar correction is used by Hayes and Norman (1984, 1994). Various adjustments to the rule are discussed by Kennedy (1998), including Tranter’s correction that accounts for fitness, fatigue, load carried, prevailing weather, and conditions under foot. Aitken (1977) has also suggested a correction for ground conditions. Ascents of the “Munros” (peaks in Scotland over 3000 ft) are assigned approximate durations by the Scottish Mountaineering Club (Bennet, 1985) that allow 4½ km per hour for distance walked and 600 m per hour for climbing ($x = 7.5$). For simplicity, orienteers use $x = 10$ (Disley, 1972).

In this paper, I first present a thorough analysis of the fell running records initially discussed by Scarf (1998). I then revisit the route choice problem considered in Figure 1. Finally, I contrast the findings with those presented recently by Norman (2004). Remaining research questions are also discussed in the final section.

**Analysis of the fell running records**

Typically, a fell race is run annually on a given course or route. For any given year, the fell racing calendar publishes the details of races including the race distance, the total climb on the race route, and the record times for men and women – these records are the fastest times recorded in the history of the race on the given route. The record times (minutes),

---

![Figure 2. The Harwich treadwheel crane. © The Harwich Society.](image-url)
distances (km), and climb (m) for each of the 300 or so races in the 1994 fell racing calendar are shown in Figure 3 (male runners). The races vary in distance from 2 to 56 km and the climb (ascent) from 90 to 2750 m. I denote the times by \( t_i \), the distances by \( x_i \), and climb by \( y_i \) (\( i = 1, \ldots, n \)). I define \( z_i = x_i + a y_i \) as the equivalent distance of race \( i \) for a given equivalence ratio \( a \). Figure 4a plots the record times against the equivalent distances for \( a = 7.92 \). Observe that the record time is not proportional to the equivalent distance; this is understandable, as runners experience greater fatigue in longer races.

It is reasonable to suppose that times vary with distance according to a power law, although other models have been proposed (Grubb, 1998). Thus in deterministic form

\[
t_i = \gamma z_i^\beta = \gamma (x_i + a y_i)^\beta. \tag{1}
\]

The parameter \( \beta \) can be interpreted as a fatigue factor; the greater \( \beta \) is than 1, the greater the fatigue with increasing distance (or equivalent distance). Runners usually refer to pace rather than mean speed, where pace = 1/(mean speed); marathon runners will talk about 5 minute-mile pace rather than a speed of 12 miles per hour, for example. Thus the pace, \( p_i \), of race \( i \) is \( p_i = t_i/\gamma z_i = \gamma (x_i + a y_i)^{\beta-1} \), and so \( \gamma \) can be interpreted as the limiting (record) pace when there is no fatigue (\( \beta = 1 \)).

Taking logarithmic transformations of time and equivalent distance (to base \( e \)), it can be seen (Figure 4b) that \( log(t_i) \) is linearly related to \( log(z_i) \), thus supporting equation (1). Scarf (1998) used an indirect method of estimation based on the fact that for a given \( a \), \( \gamma \) and \( \beta \) can be estimated by least squares. A crude search was used to determine that \( a \) which maximizes the coefficient of determination, \( R^2 \), in the linear regression of \( log(t_i) \) on \( log(z_i) \). The drawback of this method is that the standard error of \( a \) was not available. Here I consider maximum likelihood estimation for two competing models. Model 1 is the classic log-log-linear regression model with normal errors that assumes

\[
\log(t_i) = \log\gamma + \beta \log(x_i + a y_i) + \epsilon_i, \tag{2}
\]

\( \epsilon_i \sim N(0, \sigma^2), \quad (i = 1, \ldots, n). \)
This is the log-normal regression model. The model implies that

\[ E(t_i) = ye^{x_i/2}(x_i + zy_i)\rho. \]

Model 2 is an extreme value regression model (Coles, 2001) justified on the basis that the record times may be considered as minimum times from the set of past winning times for each event. Assume that

\[ -t_i \sim GEV(-u\omega, \sigma z, k), \]
\[ z_i = x_i + zy_i \quad (i = 1, \ldots, n). \]

That is, the record times follow a generalized extreme value distribution (GEV), with location and scale parameters as power-law functions of the record times is necessary for model fitting. The distribution function of the standard GEV \((u, \sigma, k)\) distribution is given by

\[ F(x) = \exp \left\{ -[1 - k(x - u)/\sigma]^{1/k} \right\}, \]

and so under model 2 the record times have a distribution function given by

\[ F(t) = 1 - \exp \left\{ -[1 + k(tz - u)/\sigma]^{1/k} \right\}. \]  (3)

It follows from (3) that

\[ E(t_i) = \left[ u - \left( 1 - \frac{\sigma}{k} \Gamma(1 + k) \right) (x_i + zy_i) \right] = \mu(x_i + zy_i), \]

as required. Here \(\Gamma(.)\) is the gamma function. Both models thus assume that the expected record times are a power-law function of equivalent distance.

The log-likelihood for the log-normal regression model is then

\[ L = -n \log \sigma - n \log \sqrt{2\pi} \]
\[ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (\log t_i - \log \gamma - \beta \log z_i)^2, \]

and the log-likelihood for the extreme value regression model is obtained from (3) as

\[ L = -n \log \sigma - \beta \sum_{i=1}^{n} \log z_i \]
\[ + \left( \frac{1}{\lambda} - 1 \right) \sum_{i=1}^{n} \log \left[ 1 + \frac{k}{\sigma} \left( t_i z^{-\beta} - u \right) \right] \]
\[ - \sum_{i=1}^{n} \left[ 1 + \frac{k}{\sigma} \left( t_i z^{-\beta} - u \right) \right]^{1/k}. \]

These log-likelihoods can be maximized using software with a standard function maximizer.

Parameter estimates for the log-normal regression model are shown in Table I, with an approximate 95% confidence interval for Naismith’s number indicated. This confidence interval is simply the parameter estimate plus and minus two standard errors. Model 1 implies that for men

\[ E(t_i) = 2.09(x_i + 8.6y_i)^{1.14} \]  (4)

and for women

\[ E(t_i) = 2.34(x_i + 10.6y_i)^{1.16}. \]  (5)

Maximum likelihood estimates for the extreme value regression model are shown in Table II, again with an approximate 95% confidence interval for Naismith’s number. The extreme value regression model implies that for men \(E(t_i) = 2.11(x_i + 8.3y_i)^{1.15}\) and for women \(E(t_i) = 2.34(x_i + 9.6y_i)^{1.17}\). Akaike information criterion (AIC) statistics (Table III) suggest that the log-normal regression model is a better fit. Thus I adopt the log-normal regression model with maximum likelihood estimates as the final model, so that expected record times are given by equations (4) and (5). Note that in all models, distance \((x)\) and climb \((y)\) are measured in kilometres and time \((t)\) is measured in minutes.

The pace model for male runners is thus \(p = 2.09(x + 8.6y)^{0.14}\). Over short races with an equivalent flat distance of 10 km, the record pace is 2.9 min per flat kilometre; for longer races with an equivalent flat distance of 30 km, it is 3.4 min per

---

Table I. Parameter estimates for the log-normal regression model (model 1) (standard errors in parentheses).

<table>
<thead>
<tr>
<th></th>
<th>Naismith’s number, (\gamma)</th>
<th>Fatigue factor, (\beta)</th>
<th>Limiting pace, (\sigma)</th>
<th>Approx. 95% CI for (\gamma)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>From Scarf (1998)</td>
<td>8.0</td>
<td>1.14</td>
<td>2.18</td>
<td></td>
</tr>
<tr>
<td>MLE</td>
<td>8.6 (0.6)</td>
<td>1.14 (0.01)</td>
<td>2.08 (0.09)</td>
<td>0.11 (0.01)</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>From Scarf (1998)</td>
<td>9.50</td>
<td>1.15</td>
<td>2.43</td>
<td></td>
</tr>
<tr>
<td>MLE</td>
<td>10.6 (0.9)</td>
<td>1.16 (0.02)</td>
<td>2.31 (0.13)</td>
<td>0.15 (0.01)</td>
</tr>
</tbody>
</table>

**Note:** Approx. 95% CI = approximate 95% confidence interval; MLE = maximum likelihood estimates.
Table II. Maximum likelihood estimates for the extreme value regression model (model 2) (standard errors in parentheses).

<table>
<thead>
<tr>
<th>Naismith's number, $x$</th>
<th>Fatigue factor, $\beta$</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$k$</th>
<th>Approx. 95% CI for $x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>8.3 (0.6)</td>
<td>1.15 (0.01)</td>
<td>2.19 (0.10)</td>
<td>0.26 (0.02)</td>
<td>0.34 (0.02)</td>
</tr>
<tr>
<td>Women</td>
<td>9.6 (0.5)</td>
<td>1.17 (0.02)</td>
<td>2.43 (0.09)</td>
<td>0.40 (0.02)</td>
<td>0.49 (0.03)</td>
</tr>
</tbody>
</table>

Note: Approx. 95% CI = approximate 95% confidence interval.

Table III. Akaike information criterion statistics for the two models.

<table>
<thead>
<tr>
<th></th>
<th>Log-normal model</th>
<th>Extreme value model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>1653.8</td>
<td>1667.8</td>
</tr>
<tr>
<td>Women</td>
<td>1911.6</td>
<td>1918.6</td>
</tr>
</tbody>
</table>

flat kilometre. For world track records ranging from 1.5 to 50 km, simple linear regressions of $\log(t)$ on $\log(x)$ give $t = 2.23x^{1.09}$ for males and $t = 2.52x^{1.07}$ for females (Scarf, 1998). These models are reasonably consistent with those established for the UK fell running records.

Data on the winning times for 38 stages of international multi-day cycling races have been analysed by Scarf and Grehan (2005). The log-normal regression model (equation 2) was fitted to these data, with fitted model $t = 1.0627(x + 8.2)^{1.0632}$. Thus $x$ is close to that obtained for Naismith’s number for runners. The authors also extended the idea of equivalent distance to account for differing terrain types by introducing an additional parameter $\delta$ that quantifies the effect of the rideability on the equivalent distance. This “terrain index” is defined so that the equivalent distance of a route of length $x$ with $y$ units of climb over terrain with terrain index $\delta$ is given by $\delta(x + y)$. $\delta = 1$ indicates perfect going (road), while $\delta > 1$ represents less than perfect going. For a journey incorporating several types of terrain, the total equivalent distance can be considered as a composite of distances over $n$ different terrain types. Then the equivalent distance for the route would be $\sum \delta_i(x_i + y_i)$, where $x_i$ and $y_i$ are the total distance and climb on terrain of type $i$ over the route, $\delta_i$ is the going index for terrain type $i$, and $x$ is Naismith’s number. An alternative formulation would be $\sum \delta_i(x_i + y_i)$, with the going index multiplicative on distance only. Further work is necessary to estimate $\delta > 1$. In the case of running or walking, the estimation of a runability index could be worthwhile.

**Route choice example**

Three routes can be distinguished for the leg from control points 1 to 2 in Figure 1: a southern route using the road down Wasdale and then turning left up Nether Beck (“road” route); a northern route following the ridge over Pillar and Scoat Fell (“Pillar” route); and a middle route via Dore Head and the southern slopes of Red Pike (“middle route”). All routes go to the saddle south of Brandreth; the Pillar and middle routes then contour and pick up the path to Black Sail Pass below Boat How crags, while the road route goes over Beck Head. The distances and equivalent distances are shown in Table IV. If $X$ is the horizontal distance and $Y$ is the climb for a route, then the equivalent distance for the route is $X + zY$, where $x$ is the equivalence ratio of climb to distance (Naismith’s number). Thus based on distance and climb alone, route A is “faster” than route B if

$$X_A + zY_A < X_B + zY_B.$$  

There appears little to choose between the middle and Pillar routes. Of course, other factors could enter the decision problem such as the runability: the Pillar route will be mostly good running on paths; the road route, although longer, is mostly very good running. Navigationally, the road route will be simpler, particularly in bad weather given the low-level running and the direct approach to the checkpoint. The runability of the middle route is likely to be the least good of the three. A variation on the middle leg that stays high beyond Gosforth Crag and contours via Scoat Tarn could be a good route. This saves significant climb and has the option of a simple approach to the control (straight on) if visibility is poor. It is not a straightforward decision about which route is best – distance, climb, runability, and navigational simplicity all impact on the choice here. However, quantification of the equivalent distance nonetheless provides useful information. Such route choice considerations could also be useful for event planners and the siting of controls may be analysed along the lines suggested by Hayes and Norman (1984).

**Discussion**

Naismith’s rule is used by walkers and runners for estimating journey times. The analysis in this paper suggests a value of 8–10 for Naismith’s number, implying that 1 m of climb is equivalent to 8–10 m
of horizontal travel. The rule is best used in its equivalence form – this then allows routes to be compared independently of fitness (flat speed). Furthermore, in its equivalence form the rule can be used to quantify accessibility; Carver and Fritz (1998) produced maps of mountainous regions with “iso-equivalent-distance” curves connecting points equally distant in time terms from access points. The rule may also be incorporated into geographical information systems software in order to calculate journey times from route profiles and used in web-based systems (Carver & Fritz, 2000).

A similar value for Naismith’s number has been obtained for cycling, which is unsurprising since the relative efficiencies of both methods of transport are similar (Howes & Fainberg, 1991). However, these results should be contrasted with the findings of Davey, Hayes and Norman (1994, 1995) and Norman (2004). Experimental results derived from running on treadmills support \( \alpha = 3.3 \); further results based on the times for a mountain road-relay support \( \alpha = 4.4 \). A possible explanation for the discrepancy between Naismith’s number for the fell running data and the number reported by Norman for the treadmill data may be that running machines overestimate the simulated gradient, given that when running “uphill” on a treadmill a runner remains stationary and so his potential energy does not change. This idea suggests that the calibration of the gradient on running machines might be done using timed experiments and a rule such as Naismith’s rule rather than using the gradient of the running surface directly.

The discrepancy between Naismith’s number for the fell running data and the number reported by Norman for the mountain road-relay data \( \alpha = 4.4 \) is more difficult to explain. However, a contributing factor may be that in fell races steeper paths will in general be in poorer condition than more moderately inclined paths. Consequently, steeper paths may be slower not just as result of the extra climb but also because of the reduced runability – if this is indeed the case, then in fell races climb or ascent will have a greater slowing effect on running speeds and hence times than in road races. The correction described by Hayes and Norman (1984) would go some way to explain this discrepancy. Also, the work of Minetti (1995) on the metabolic cost of human locomotion on gradients argues that on very steep routes runners are moving less efficiently than on more moderate slopes. This again suggests a greater slowing effect of ascent and descent “off-road”.

It should be emphasized that it is not so much the complete rule but the value of Naismith’s number that is important for individual runners. This is because individuals will differ in their fitness and so a person can use, in time calculations, their own personal value for their pace over good, flat terrain. One need only calculate the equivalent distance of a route and then multiply by one’s own personal pace measure. The notion of equivalent distance is also intuitive: one can ask if going round is equivalent to going over (a hill, say). Of course, other factors will impact on the decision, including the going under foot (runability) and the navigational demands of route.

Although I have not explicitly discussed descent, this is accounted for in the equivalent distance calculations considered here. In the analysis of the fell running records, the use of Naismith’s number assumes that all journeys start and finish at the same height – fell races typically start and finish in the same place. Consideration of journeys that are pure ascents is a different problem. In a simple approach to route choice comparison based on Naismith’s rule without Langmuir’s correction, we do not need to consider descent. This is because where two routes between two points differ in the amount of distance and climb, on the route with extra climb the amount of extra climb will be equal to the amount of extra descent on that route – this is a logical consequence of the fact that the routes have common starts and finishes. Thus, necessarily in a route choice problem, “what goes up, must come down”! Where ascents or descents are very steep, then more complex rules will be appropriate.

Finally, the support that the analysis of fell running records lends to Naismith’s rule suggests that there is now some justification in describing the idea of Naismith as a law, or a rule. The extensions to Naismith’s rule that have been proposed by others such as Langmuir and Tranter remain theories and it would be interesting to test them. Furthermore, an investigation of veteran fell running records might shed some light on the question of whether Naismith’s number depends on age. Also, analysis of complete race results rather than just record times may be useful for determining a fitness effect. More extensive data on hill walking similar to those analysed by Kennedy (1998) would be necessary to confirm that the results obtained in this paper are applicable to walkers.

<table>
<thead>
<tr>
<th>Route</th>
<th>Distance (km)</th>
<th>Climb (m)</th>
<th>Equivalent distance ((\alpha = 10))</th>
<th>Equivalent distance ((\alpha = 8))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Road</td>
<td>12.3</td>
<td>700</td>
<td>19.3</td>
<td>17.9</td>
</tr>
<tr>
<td>Middle</td>
<td>10.0</td>
<td>800</td>
<td>18.0</td>
<td>16.4</td>
</tr>
<tr>
<td>Pillar</td>
<td>10.5</td>
<td>750</td>
<td>18.0</td>
<td>16.5</td>
</tr>
</tbody>
</table>
Acknowledgements

I am grateful to the referees for their comments and suggestions, which greatly improved the paper. Thanks also go to Susan Harvey of HARVEY Maps for the digital reproduction of the section of the Lake District map used in Figure 1.

References


