An analysis of probability of area techniques for missing persons in Yosemite National Park

Article *in* Applied Geography · February 2014

DOI: 10.1016/j.apgeog.2013.11.001

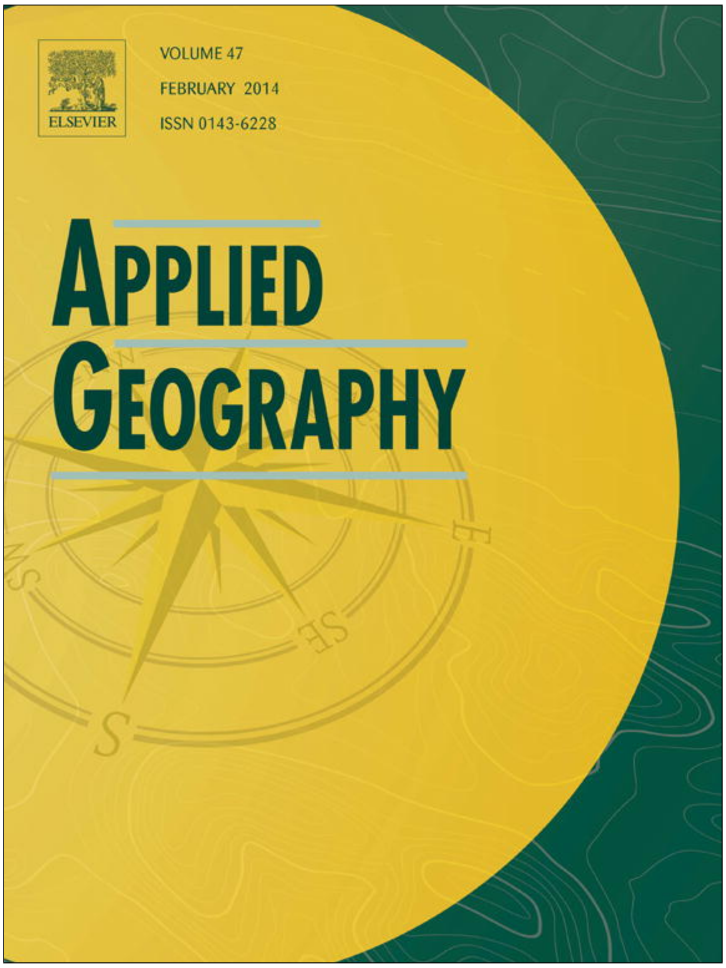
4 authors:

Paul Doherty
U. S. Department of Homeland Security

Qinghua Guo
Peking University

Jared Doke
University of Kansas

Don Ferguson
West Virginia University


ELSEVIER
VOLUME 47
FEBRUARY 2014
ISSN 0143-6228
APPLIED GEOGRAPHY

Applied Geography 47 (2014) 99–110

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog

An analysis of probability of area techniques for missing persons in Yosemite National Park

Paul J. Doherty [a], Quinghua Guo [a], Jared Doke [b], Don Ferguson [c,*]

[a] School of Engineering, Sierra Nevada Research Institute, University of California, Merced, CA 95344, USA
[b] Department of Geography, University of Kansas, 213 Lindley Hall, 1475 Jayhawk Blvd, Lawrence, KS 66045-7613, USA
[c] Mountaineer Area Rescue Group, PO Box 204, Kingwood, WV 26537-0204, USA

Keywords:
Search theory
Time geography
Travel cost
Wildland search and rescue

A B S T R A C T

Study of wilderness search and rescue (WiSAR) incidents suggests a dependency on demographics as well as physical geography in relation to decisions made before/after becoming lost and subsequent locations in which subjects are found. Thus an understanding of the complex relationship between demographics and physical geography could enhance the responders' ability to locate the subject in a timely manner. Various global datasets have been organized to provide general distance and feature based geostatistical methods for describing this relationship. However, there is some question as to the applicability of these generalized datasets to local incidents that are dominated by a specific physical geography. This study consists of two primary objectives related to the allocation of geographic probability intended to manage the overall size of the search area. The first objective considers the applicability of a global dataset of lost person incidents to a localized environment with limited geographic diversity. This is followed by a comparison between a commonly used Euclidean distance statistic and an alternative travel-cost model that accounts for the influence of anthropogenic and landscape features on subject mobility and travel time. In both instances, lost person incident data from years 2000 to 2010 for Yosemite National Park is used and compared to a large pool of internationally compiled cases consisting of similar subject profiles.

Introduction

Wilderness search and rescue (WiSAR) consists of four basic processes that are defined as: Locate, Access, Stabilize and Transport. Each of these elements presents a unique geographic problem which provides a novel and largely unexplored testing ground for the spatial sciences. Of these elements, "Locating" the lost subject(s) can often prove particularly challenging as a result of the potential interaction between time and geography as defined by Winter and Yin (2010). Stated more succinctly, in order to find a lost subject the responder must overlap the subject in both time and space. This issue is further exacerbated by the often disproportionately limited number of responders for the large geographical area that is defined as the search area. The need to search such large areas often times contributes to the delay in locating the lost subject and subsequently impacts their chances of survival. In fact, if missing subject is not found within the first 51-h of a search, the chances of surviving decline significantly (Adams et al., 2007).

Several options are available to the responder for improving their response capability including: 1) increasing the number of responders available to search, 2) enhancing the available responders' Probability of Detection (POD – likelihood of detecting the subject if they were present) or 3) reducing the size of the search area through improved assignment of geographic Probability of Area (POA – likelihood of the search being present in a subregion within the search area) (Cooper, Frost, & Robe, 2003). Limitations on personnel and capabilities forces attention on to methods that can be used to better define POA.

Demographics and the environment have a strong influence on where lost persons are ultimately found, and the decision that are made by the subject(s) both before and after becoming lost. The demographics, or category, of a lost subject includes the activity in which they were involved in prior to becoming lost, their age and their medical condition particularly as it relates to their cognitive ability. These factors greatly influence the relationship between the lost subject and their environment. For example, individuals characterized as "Hikers" are more prone to rely on linear features such as roads and trails for travel. These individuals often become lost as

* Corresponding author. Tel.: +1 304 290 9118.
E-mail addresses: pjdohert@gmail.com (P.J. Doherty), qguo@ucmerced.edu (Q. Guo), jared.doke@gmail.com (J. Doke), dferguso@mix.wvu.edu (D. Ferguson).

0143-6228/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.apgeog.2013.11.001

a result of active or passive decisions made in relation to these travel aides (ie decision to leave the travel aide for some purpose or confusion along the feature). And when lost these individuals tend to focus locating some form of linear feature to assist with re-orientation (Koester, 2008; Syrotuck, 2000). Both before and after becoming lost, individuals are forced to make decisions based on their immediate environment. This would suggest that geographical knowledge of the search area could enhance responders' ability to located lost persons or at least more accurately assign POA to regions within the search area.

To assist the responder in better understanding complex interactions between subject and environment, several global datasets have been compiled that span a variety of subject categories (Hill, 1999; Koester, 2008; Syrotuck, 2000). Currently the largest of these datasets, consisting of over 50,000 incidents from around the world is the International Search and Rescue Database (ISRID - Koester, 2008). ISRID contains forty-one different subject categories with sub-divisions based on Bailey's terrestrial eco-region domains (Bailey & Ropes, 1998 – Fig. 1) and population density of the location of the Initial Planning Point (IPP). The IPP is typically defined as either the Point Last Seen (PLS) or the Last Known Point (LKP) (Ferguson, 2008; Theodore, 2009). Utilizing the categorized datasets within ISRID basic geostatistical analyses were performed to consider such parameters as distance from IPP to Find Location (Euclidean Distance – ring model), geographical description of the find location, difference in elevation between the IPP and Find Location and distance from nearest linear feature (road, trail, drainage) to Find Location. Additionally post-search interviews of search subjects were conducted to obtain an estimate of the time the subject was mobile and how far the subject had deviated from their intended destination. Cumulative information was used by Koester to develop a generic lost person behavioral profile based on subject category (Koester, 2008).

As the ISRID database is composed of world-wide incident data, there is some concern as to its accuracy for local incidents that occur within an area with limited geographic diversity. This study provides a comparison of search incident data for the "Hiker" category from ISRID ($n = 568$) and Yosemite National Park for years 2000–2010 ($n = 130$). The ring model (Euclidean Distance from IPP) is often used to assist in assigning probability to regions within the search area. This metric provides an easy to use format for comparing the global and local datasets. Additionally, comparisons are made between the ring model and a mobility (cost-distance) model that accounts for the influence of terrain, vegetation, travel

Fig. 1. Bailey Eco-Region Domains used in classifying the search area within the International Search and Rescue Database (Bailey & Ropes, 1998; Koester, 2008).

aides and travel barriers. Arguments are made suggesting that the mobility model provides a more realistic interpretation of lost person behavior compared to simple Euclidean distance.

Probability of area

It is common for WiSAR incidents that extend beyond the first operational period to encompass large geographical areas. Assuming a nominal walking speed of 5.1 km/h (Browning et al., 2005), the search area could theoretical expand to 82 km^2 (20,200 acres) in just 1 h not knowing which direction the subject may have traveled from the IPP (Fig. 2). In order to effectively manage such operations, responders must establish a method of prioritizing the search area, often times utilizing some interpretation of lost person behavior (Koester, 2008). This typically leads to the establishment of region of probability suggesting the most likely locations of where the subject is to be found. These concepts are grounded in Operations Research that were initially developed as maritime and aviation search techniques (Cooper et al., 2005; Koester, Cooper, Frost, & Robe, 2004; Koopman, 1980; Stone, 1989, 2007). The formal study of "Search Theory" as a discipline grew out of military operations conducted during World War II. The mathematical driven process of how to search for missing, lost, and hidden objects was used to locate enemy submarines as well as to recover lost allied ships and downed pilots (Frost, 1999). These initial concepts have been adapted to the nuances of ground-based search operations by Cooper (Cooper et al., 2005) and others, and has been adopted into the National SAR Plan within the United States (Land SAR Addendum, 2011).

In WiSAR, grid-based searches are often not possible due to terrain features, large geographical areas or the limited number of resources available. Based on this, the first step in allocating resources to find these missing persons is to begin searching the immediate vicinity of the IPP which is typically either the Point Last Seen (PLS) or Last Known Point (LKP) based on the availability of substantial evidence. These strategies assist in defining the general extent of the search area which then must be divided into feature-based polygons known as Planning Regions where experts use their collective knowledge to prioritize resource allocation. Search planners utilize a combination of knowledge on Lost Person Behavior (LPB), investigative methods, consensus techniques (Bownds, Ebersole, Lovelock, O'Connor, & Toman, 2007; Mattson, 1976) and terrain analysis to assign quantitative values to Regions based on the probability that it contains the subject (aka Probability of Area [POA], Cooper, 2005). These Regions are than sub-divided into Segments that can be assigned to Search Teams, Fig. 3. An important distinction is that the boundaries of the Planning Regions are defined entirely on anticipated subject behavior while the Search Segments boundaries are defined based on how the area is to be searched. For example, in establishing a manageable size Segment easy to recognize terrain features such as drainages or significant changes in slope along with shifts in land cover are often used as boundaries for but these features may not impact the size or shape of the Region. Thus geographic information plays a critical role in both estimating the most probable location of the subject as developing tactics for effectively searching an area.

The search planning process incorporates several probabilistic concepts: Probability of Area (POA), Probability of Detection (POD), and Probability of Success (POS) as described in equation (1) below (NSARC, 2011)

$$POS = POA \times POD \tag{1}$$

The Probability of Success (POS) in search theory is completely dependent upon the boundary of the area (polygon) being searched

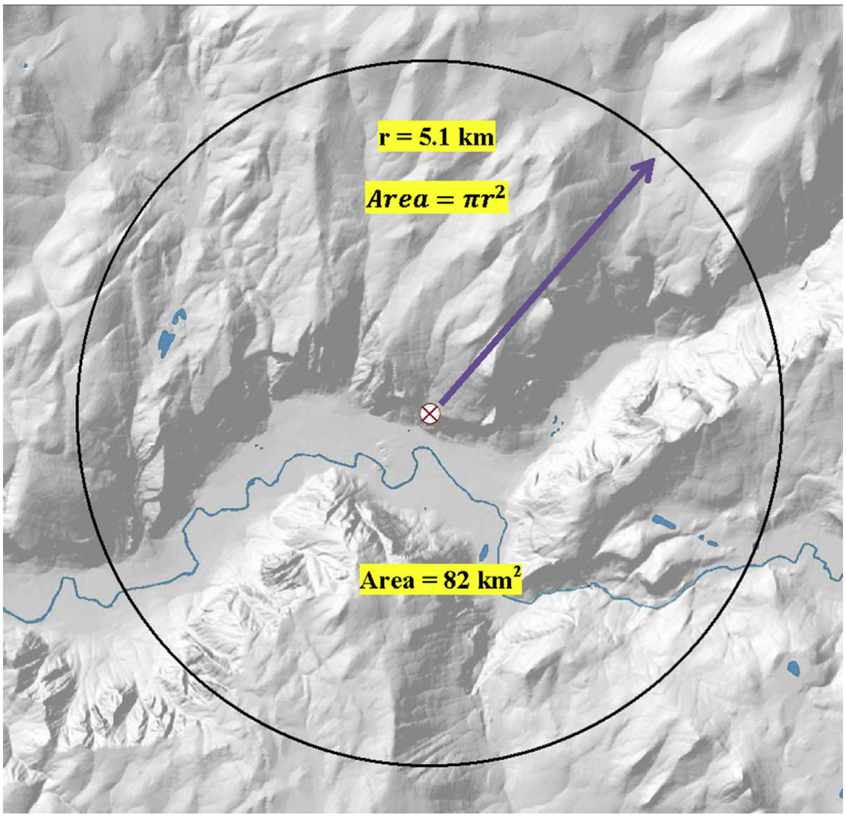


Fig. 2. Assuming a nominal walking speed of 5.1 km/h produces a search area that could theoretical expand to 82 km^2 (20,200 acres) in just 1 h not knowing which direction the subject may have traveled from the IPP.

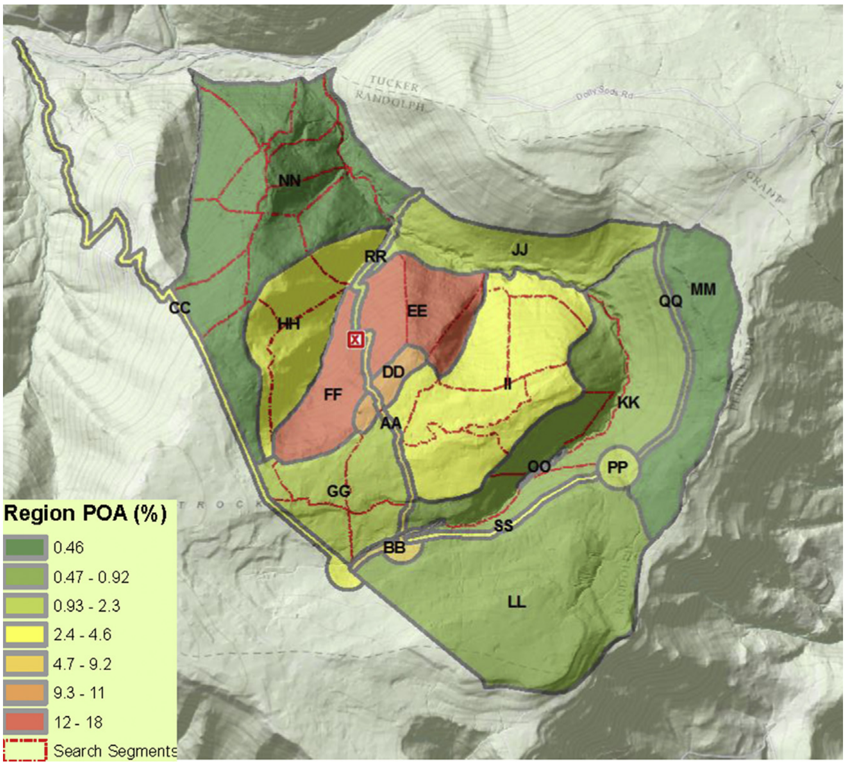


Fig. 3. Example of Probability Regions color coded based on the likelihood of it containing the subject based on subjective analyses. Search Segments (dashed border) are defined so boundaries are easily identifiable by teams in the field. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

actually containing the missing person (known as POA) and an accurate assessment of how well a search area was covered by a team (Probability of Detection [POD]). POD can be explained using the following equation.

$$POD = 1 - e^{-c} \tag{2}$$

where c is the coverage and e is the base of the natural logarithm. Coverage is the ratio of two areas: search effort as track length multiplied by width (the polyline representing searchers movements and the buffer indicating where they can effectively search), divided by the total area searched. In search planning like most complex decision making processes, errors in judgment (underestimated POA, overestimated POD) early in the planning stages significantly hampers the search effort despite subsequent decisions. The objective of any search operation is to maximize POS as quickly as possible by increasing POA and POD. At this time, most WiSAR operations utilize these concepts in a variety of ways but with little geospatial resources to accurately measure either POA or POD. With the integration of Global Position Systems (GPS) for searchers, a quantitative index of POD can be obtained by measuring coverage based on searchers' GPS receiver track-log, but the POA is still very much subjective.

Several authors have compiled results from the analysis of historical search incidents the intent of providing the responder a basic understanding of lost person behavior to facilitate the prioritization of the search area (Koester & Stooksbury, 1995; Syrotuck, 2000). A systematic collection of information from missing person incidents was conducted by Koester in the formation of the International Search and Rescue Database (ISRID; Koester). Analysis of the nearly 50,000 incidents contained within ISRID resulted in a text by the same author (Koester, 2008). In summarizing the findings, the text categorizes missing person by demographics (e.g. age, gender, mental status), participating activity (e.g. hiking, hunting, skiing) and environment (eco-region and generalized terrain). Basic statistical analyses provides responders with a variety of parameters that can be used to assist in prioritizing the search area, and includes: measured dispersion from intended direction of travel, elevation change between IPP and Find location, distance subjects were found from travel aides such as roads/trails, and a feature type for the find location (e.g. structure, field, trail). While comprehensive, the analysis provided by Koester has several limitations, such as a lack of discrimination between subjects categorized as lost versus those that suffered from a medical condition resulting in rescue. Additionally, given the coarseness of the geographical references (eco-region and generalized terrain: Mountainous versus Flat) studies have shown some discrepancy between ISRID's global average findings and those from a localized dataset (Doke, 2012).

Ring model

The ring model, amassed from a categorized statistical analysis of past incidents, consists of a series of concentric rings originating from the IPP that represent the quartile and 95% Euclidean distances to the find locations (Syrotuck, 2000). An example of the ring

Table 1
Comparison of Euclidean distances traveled from the IPP to the found location for hikers.

	Yosemite (km)	ISRID (km)
n	130	568
25%	1.1	1.1
50%	1.8	3.1
75%	4.0	5.8
95%	16.9	18.3

model using the ISRID data for the "Hiker" subject category is shown in tabular form in Table 1 and graphically in Fig. 4.

Mobility model

Mobility is defined in ISRID as "the amount of time the subject was moving", and is based primarily on post-search interviews with the subject. Thus the reported values are highly subjective and limited to those cases in which post-search interviews were possible (i.e. non-fatality). Table 2 recreates the Mobility data as reported in Koester's text on LPB for the Hiker category in a "Temperate" environment (Koester, 2008).

Mobility models have been proposed as a POA method, however to date, little GIS integration has been employed in this process (for exceptions see Filipkowska, Koester, Chrustek, & Zaród, 2012; Magyari-Sáska & Dombay, 2012). The mobility models attempt to estimate how far the subject may have traveled following a path in any direction from the IPP. Unlike the previously defined area as shown in Fig. 2, in reality it is understood that subject travel is hampered and aided by various terrain features and vegetation, thus the distance traveled in a given amount of time would be dependent on the path followed. This approach is equivalent to cost-distance which has been used extensively for infrastructure development (Van den Broek et al., 2010), wildlife habitat analysis (Nikolakaki, 2004), and anthropological studies (Whitley & Hicks, 2003). In these models, costs are calculated by applying a least-cost path algorithm to a speed raster and a resistance raster (Adriaensen et al., 2003). Tobler (1965) was the first to use the Imhof (1950) "hiking function" (equation (3)) to calculate the cost associated with traversing a landscape with slope. The "hiking function" estimates the velocity of travel for hikers across different slopes assuming a nominal walking speed of 5.0 km/h. According to Tobler (1991) in his overview of non-isotropic modeling between pedestrian movement and slope, travel speed can be estimated as:

Table 2
Comparison of mobility values for hikers in Yosemite derived from isochrones versus a global dataset (ISRID) based on reported mobility.

	Yosemite (h)	ISRID (h)
n	130	232
25%	0.5	0.0
50%	1.0	3.0
75%	1.5	6.0
95%	6.0	14

$$v_s = 6\,e^{-3.5\mathrm{abs}(\mathrm{Slope}+0.05)} \tag{3}$$

where e is the natural logarithm and v_s is pedestrian velocity defined by a mathematical function based on the slope of terrain in degrees. This yields a function that is asymmetric about zero slopes because it is generally faster to travel down hills than up hills. Naismith's Rule with an additional correction developed by Langmuir offers an alternative algorithm for predicting hiking times in mountainous terrain (Langmuir, 1984). The corrected algorithm assumes a nominal walking speed of 4.0 km/h with 30 min added for every 600 m of ascent. Descent rates are also impact by subtracting 10 min for every 300 m of descent for slopes between 0° and 12°, and adding 10 min when the slope is greater than 12°.

Often times it is not the hike speeds or travel times that are determined but energy expended to cross the terrain. Minetti et al. considered the energy costs of walking and running across various slopes (Minetti, Moia, Roi, Susta, & Ferretti, 2002) and found a preferred walking speed (4.0 kph on level ground) minimum energy output occurred at a slope of approximately −6°. The empirical results from Minetti et al. were extrapolated to a slope of 60° as a cost-distance model was used to suggest accessibility bias in ecology sampling (Jobe & White, 2009). However, the use of energy expenditure as a metric for modeling lost person behavior may not be appropriate. While habitual travelers in a given area may

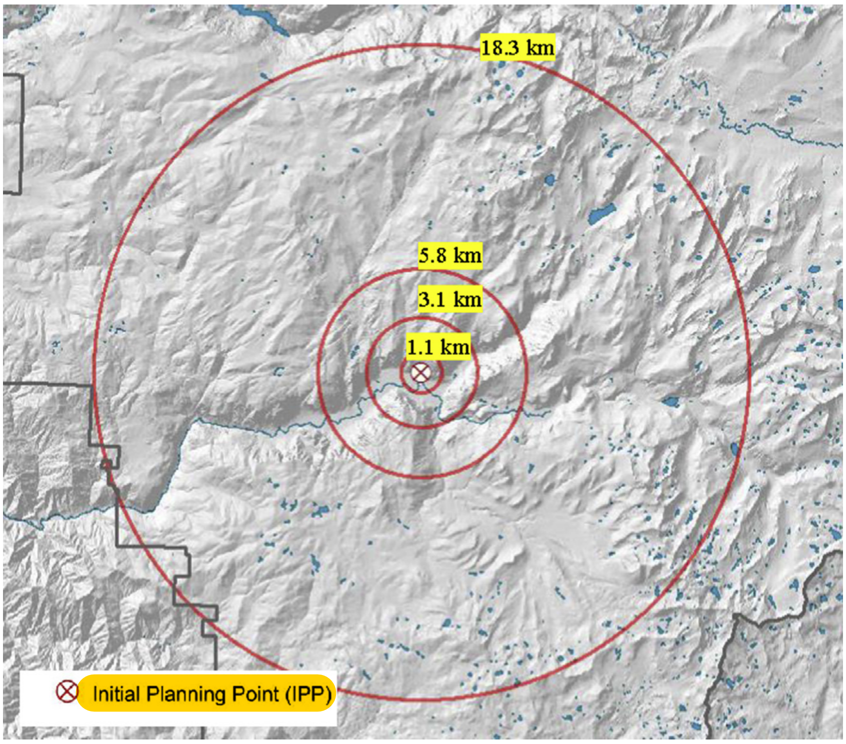


Fig. 4. Rings based on Euclidean distance (crow's flight distance) are often plotted from the Initial Planning Point (IPP). These distances correspond to the lower quartile, median, upper quartile, and 95th percentile of distance (measured in ether miles or kilometers) of lost subjects collected by the ISRID.

attempt to identify the path of least resistance and thus lowest energy output, the lost person unfamiliar with their surroundings only knows their immediate environment, and thus could not be expected to necessarily follow the path of least resistance. This would seem to support the use of a cost-distance model to predict distance traveled in a given amount of time assuming a known speed along level, unimpeded terrain. An earlier study conducted by Ralston (1958) considered energy-speed relation during level walking and provided a method for converting energy expended to walking speed given an assumed level walking speed. Utilizing this relationship the functions provided by Jobe–White are compared to Tobler's Function and Naismith's Rule shown in Fig. 5. Results for the Jobe–White and Naismith's Rule are shown with their derived level ground preferred walking speeds (4.0 km/h) and scaled to Tobler's preferred walking speed of 5.0 km/h.

The functions provided above do not account for the effects of land cover, the availability of travel aides or the presence of travel barriers. Anecdotally one could anticipate that regardless of slope, travel through a heavily vegetated area could be considerably slower than along a road or pathway. The effect of land cover on off-trail hiking velocity has been estimated by Imhof (1950) to be a reduction factor of 0.6*x*. Other studies have provided a more detailed classification of the impact of off-trail travel but none have attempted to utilize a standard definition land cover (Demczuk, 1998; Jobe & White, 2009; Soule & Goldman, 1972). Table 3 outlines three classifications found in literature.

Land cover impedance can be described as the area between the curve for Tobler's on and off-trail functions. This calibration of the impedance grid will help avoid using crude oversimplifications for slope and land cover impedance that often reduces the usefulness of such models (Bateman, Garrod, Brainard, & Lovett, 1996). If these methods represent actual pedestrian capabilities the impedance raster grid can then be used for travel-cost simulation where distance traveled is a function of estimated travel time (time since last seen), the coordinates of the IPP, and the impedance values of surrounding raster grids. The product would be a minimum potential path area.

Materials and methods

Study area

Yosemite National Park encompasses nearly 1882 km^2 and is located on the western slope of the Sierra Nevada mountain range

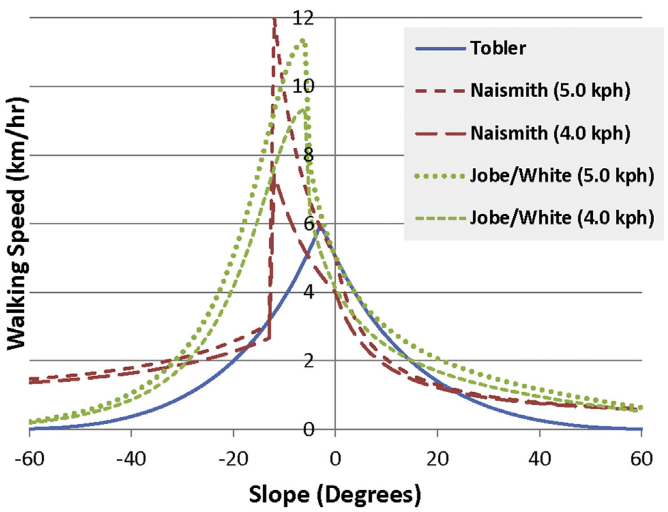


Fig. 5. Comparison of the various slope speed based functions for human foot travel at preferred walking speeds.

Table 3
Comparison of energetic cost coefficients for different terrains.

Land cover	Soule–Goldman	Demczuk	Jobe–White
Road (blacktop)	1.0	1.0	1.0
Dirt (trail – hardpack)	1.36		1.2
Light brush	1.64		1.31
Heavy brush	1.81	1.8	1.59
Swamp	2.13		1.87
Sand	2.5		
Stream			Function of stream order

in central California. To the west is the San Joaquin Valley and to the east is the Great Basin. It is fairly accessible as it lies 241 km east of San Francisco and about 482 miles north of Los Angeles (Fig. 6). Although, there are approximately 344 km of paved roads serviced by five separate entrance stations, more than 95% of the park, or roughly 1771 km^2, is designated as wilderness. About 1287 km of trails cut through this wilderness, including the Pacific Crest Trail and the John Muir Trail, within the park (National Park Service, 2011).

The park's physical landscape is a product of glaciation, and includes granite cliffs, deep narrow canyons, and a U-shaped valley that has become the focal point of the park. Elevation ranges from about 2000 ft to 13,123 ft. Yosemite is characterized by a Mediterranean climate with long hot summers and mild winters, with a Bailey's Eco-Region Domain of Temperate – Mountainous extending almost the entire region. Precipitation can vary from 36 inches at 4000 ft to 50 inches at 8600 ft. This climate, coupled with the variations in elevation, creates five major vegetation zones: chaparral/oak woodland, lower montane, upper montane, subalpine, and alpine (National Park Service, 2011).

Georeferencing

The data obtained for this analysis are derived from Yosemite National Park's Search and Rescue Case Incident reports for the years 2000–2010. Access to these reports was granted by the National Park Service Division of Visitor Protection (Permit 1024-0236). During this eleven year span, Yosemite National Park responded to 2308 total SAR Incidents. This includes both genuine searches for lost people as well as rescues in which the actual location of the individual was known. Of these SAR incidents, 2201 incident reports were available for review. Out of these reports, 393 true search incidents were identified and considered for this study; the remainder were excluded because the location of the victim was known from the beginning . Each search incident was critically analyzed and must meet all of the following criteria in order to be retained for the current study:

1. The incident must have been a ground-based search incident. Searches for downed aircraft or searches involving bodies of water were not included.
2. The incident must have had a distinct PLS/LKP/IPP that can be georeferenced within the Yosemite National Park boundary.
3. The incident must have a distinct found location that is georeferenced within, or within walking distance of, the Yosemite National Park boundary.
4. An official SAR response must have been initiated by the National Park Service.

A total of 213 search incidents met these criteria and incidents were georeferenced from text based information as part of GIS research described in Doherty, Guo, Liu, Wieczorek, and Doke

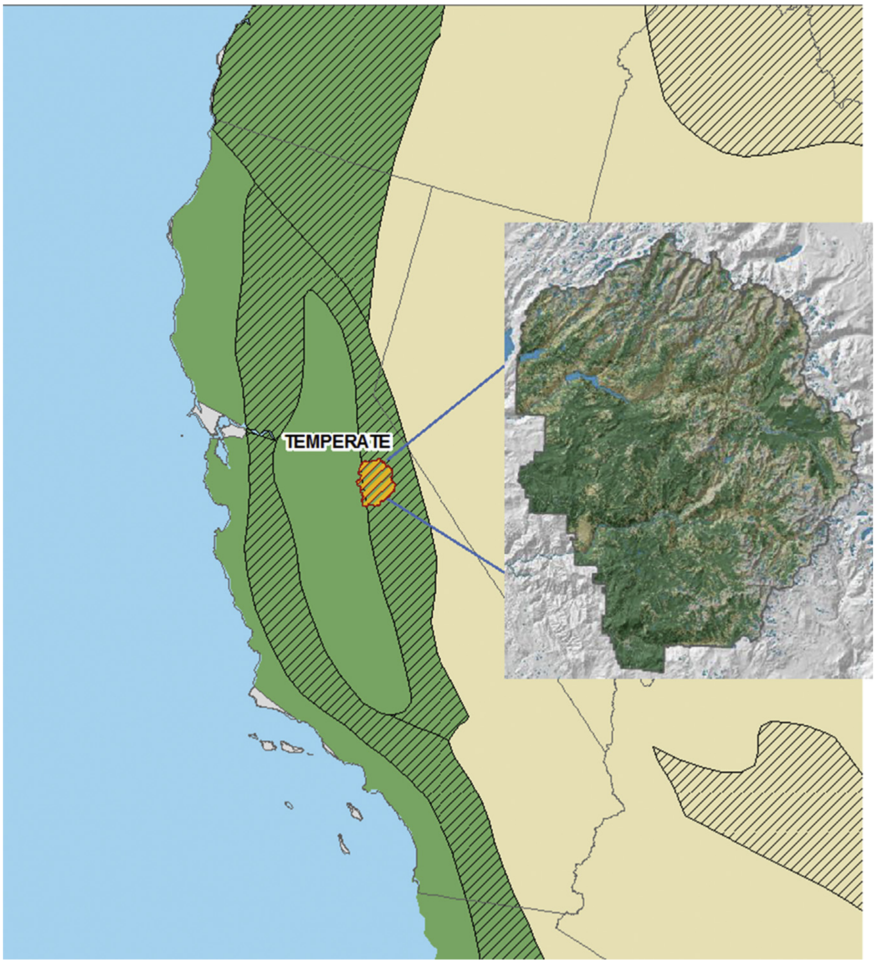


Fig. 6. Yosemite National Park.

(2011) using the point-radius method (Wieczorek, Guo, & Hijmans, 2004).

Ring model

Descriptions and details provided in the incident reports were used to georeference locations for the IPP and Find creating an incident data pair. The Euclidean distance (*D*) for each data pair of all 130 Hiker incidents was determined and provided a measure of the lower quartile, median, upper quartile, and 95th percentile. This observed distribution of *D* for Yosemite was then compared with the expected distribution of the ISRID data ($n = 538$) using a Chi-square Goodness of Fit Test with a significance level of 0.05. Standard geoprocessing and calculation of *D* was done using ArcGIS 10.1 (Esri, 2012).

Mobility model

The mobility model used in this study is similar to the Travel Time Cost Surface Model (TTCSM) used by the National Park Service (Sherrill, Frakes, & Schupbach, 2010). The model provides an estimate of travel time using readily available geospatial products such as road, trail, and stream networks (Stream Order), digital elevation models and land cover data (National Land Cover Dataset). The model consists of two basic components: speed surface and cost surface. First a simple speed surface is created based on Tobler's Hiking Function as is defined earlier (using equation (3)) with slope derived from a 10 m digital elevation model of the park. Slopes greater than 60° were consider impassable by conventional foot traffic. This speed surface does not account for energy expenditure and assumes the subject would be able to maintain a nominal speed across a flat surface (slope = 0°) of 5.0 km per hour (kph).

The cost surface is defined as a function of the impedance to foot traffic that would be imposed by the presence of various geographical features compared to a nominal paved surface such as a side walk or roadway. Impedance values range from 0 to 100%, with 0% being no impedance (e.g. paved roadway or a well-maintained trail) and 100% being absolute impedance (e.g. high fenceline or large body of water). A value of 25 would represent a feature that was 25% slower to cross (e.g. pasture/field) than the nominal surface (paved roadway).

The impedance raster is formulated using both vector and raster data layers. Vector layers include roads, trails, utility right-of-ways, fencelines (actual and virtual), streams and bodies of water utilizing data acquired from the NPS Integrated Resource Management Applications (IRMA) Data Portal (NPS, 2013). While maintained roadways offered 0% impedance, the maintenance level of the trail as recorded in the feature attribute table was used to assign impedance values with poorly maintained or unmaintained trails given a higher impedance value, Table 4.

The ability to cross streams will depend on its size and flow accumulation. First order streams that at higher elevations are rarely impossible to cross. However, at lower elevations the flow accumulation increases and streams become more difficult to cross

(Jobe & White, 2009). To account for this variation, stream impedance was assigned based on Strahler Stream Order. A combination of hydrological vector data and DEM derived features are used to represent streams and provide an estimate of stream order (Strahler) using the Stream Order function within the Spatial Analyst Toolbox in ArcGIS 10.1. The higher the Stream Order the greater the impedance, Table 5. Large bodies of water (lakes, ponds, etc) were considered absolute barriers since only foot traffic was being considered.

As noted above, previous studies did not consider the use of standardized definitions of land cover in considering their impact on a cost-distance model. The underlying land cover for this study was derived from the National Land Cover Dataset (Fry et al., 2011) with features being reclassified similar to costs defined in Sherrill (Sherrill et al. classify cost with respect to Percent of Maximum Travel Speed – PMTS, where the model used in this study is defined as impedance, 1/PMTS) (Table 6).

Once all vector layers were reclassified, they were combined using a hierarchical overlay method that prioritizes layers based on their importance. The final impedance raster is created using a hierarchical overlay method where factors are listed in decreasing priority: fenceline, road, trail, impassable slope, bodies of water, streams (Strahler order), utility right-of-ways and land cover. This impedance (cost surface) layer is then combined with the speed layer to obtain an estimated travel speed surface (km per hour) for the area of interest with foot travel speeds ranging from slightly above the nominal speed (maximum speed at a slope of −0.05) to zero. Because the Path Distance Tool within ArcGIS 10.1 assumes the cost surface defines ease of crossing a cell as opposed to the impedance represented by the cell, the inverse of the cost surface as defined above is used as input. Starting at the IPP, the accumulated cost to travel across the cost surface is evaluated in an anisotropic manner in order to account for traveling either up or downslope in moving away from the IPP. The Path Distance tool within ArcGIS provides a means of considering travel in all directions producing a travel time surface suggesting the minimal time required to reach a destination starting from the IPP.

Based on the algorithm discussed above, a Cost Surface for all of Yosemite National Park was generated and used to produce Path Distance surfaces from the IPP of all 130 missing Hiker incidents with isochrones representing distance traveled in a specified amount of time, Fig. 7.

Subject mobility was determined by sampling the Path Distance surface isochrones at the Find location providing an estimate for the minimum amount of time (T_{min}) for the subject to reach the Find location from the IPP based on the Path Distance Cost Surface. The model is not intended to take into account the fact that a person may wander, stay in one place, leave an area and return, or any impact from weather but rather it is an empirical model based on physical limitations imposed by terrain and environmental features. Nor does it account for energy expenditure, as previously noted, which could potentially increase T_{min}. We summarized hiker mobility statistics in a format that could be directly compared to the ISRID as defined by Koester (2008). The lower quartile, median, upper quartile, and 95th percentile were calculated for each for the hiker category and T_{min} values are compared to the global dataset compiled in the ISRID using a Chi-square Goodness of Fit Test. We statistically compared the ring and mobility models using bivariate correlation of D and T_{min} for each hiker search incident, Fig. 8.

Table 4
Assigned impedance based on trail maintenance level.

Trail_Class	Description	Walking impedance
1	Minimal/undeveloped trail	25
2	Simple/minor development trail	20
3	Developed/improved trail	15
4	Highly developed trail	5
5	Fully developed trail	0

Table 5
Assigned impedance based on stream size and difficulty crossing (Strahler stream order).

Strahler stream order	Impedance
1	30
2	40
3	50
4	70
5	80
6	90
7	99

Results and discussion

Ring model

Of the 213 incidents georeferenced from the Yosemite data, 130 were cases involving lost hikers were included in the analysis. Overall, the Euclidean distance (D) from each georeferenced IPP to its corresponding georeferenced found location was calculated ($n = 130$, mean = 3.34 km ± 0.25). When the local Yosemite hiker sample was compared to the global ISRID data (considering only the data from Temperate, Mountainous Eco-Region), there were statistically significant differences (Table 1; $n = 130$, $\chi^2 = 15.4$, $P < 0.01$). This means that the distances that hikers travel, as the crow flies, from the IPP to the found location in Yosemite was significantly different than the same corresponding distances that hikers travel at the global scale. The values in the first quartile (25%) were equal, but were greater in the ISRID for 50%, 75%, and 95% values.

Table 6
Assigned impedance value based on land cover classification code.

LCCC	Description	Impedance
11	Open water	99
12	Perennial ice/snow	80
21	Developed, open space	20
22	Developed, low intensity	20
23	Developed, medium intensity	30
24	Developed, high intensity	40
31	Barren land (rock/sand/clay)	60
32	Unconsolidated Shore	70
41	Deciduous forest	50
42	Evergreen forest	50
43	Mixed forest	50
51	Dwarf scrub	75
52	Shrub/scrub	75
71	Grassland/herbaceous	50
72	Sedge/herbaceous	50
73	Lichens	20
74	Moss	20
81	Pasture/hay	20
82	Cultivated crops	30
90	Woody wetlands	80
91	Palustrine forested wetland	80
92	Palustrine scrub/shrub wetland	80
93	Estuarine forested wetland	80
94	Estuarine scrub/shrub wetland	80
95	Emergent herbaceous wetlands	80
96	Palustrine emergent wetland (persistent)	80
97	Estuarine emergent wetland	80
98	Palustrine aquatic bed	99
99	Estuarine aquatic bed	99

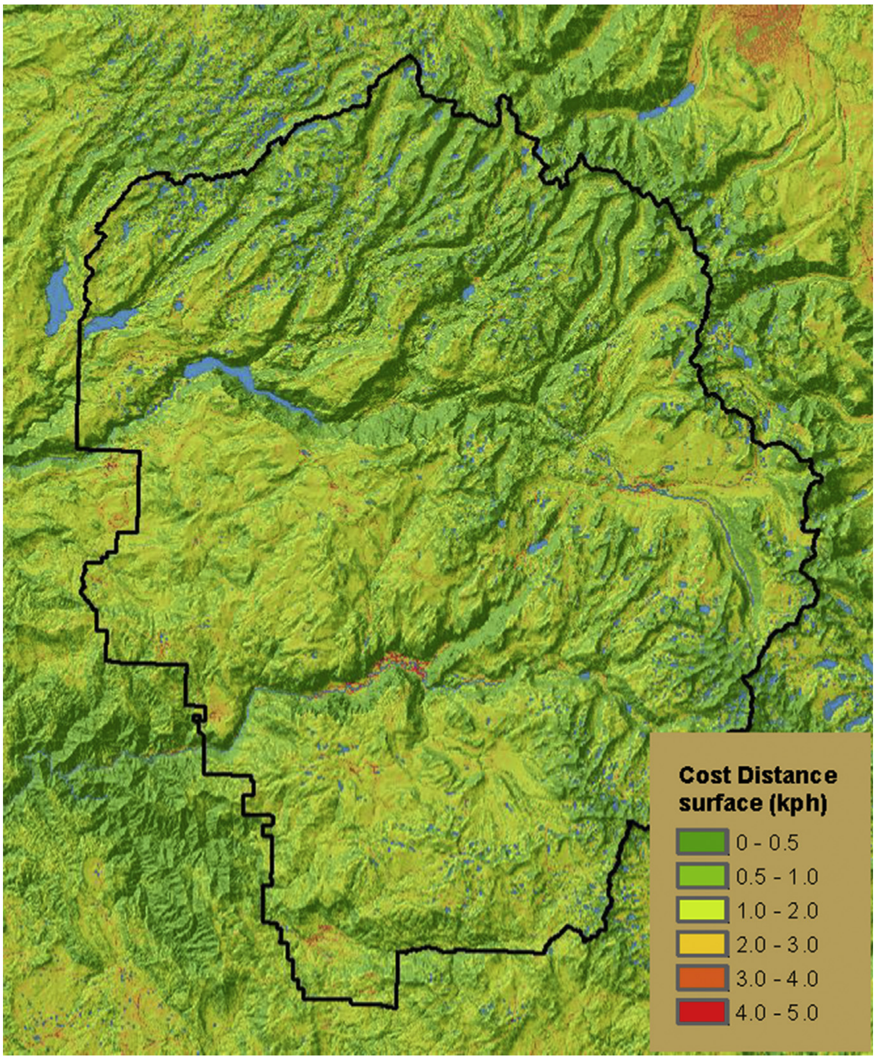


Fig. 7. Cost-distance surface for Yosemite National Park based on Tobler Hiking Function, impassable slope (<60°), availability of travel aides, land cover and presence of travel barriers assuming a nominal walking speed of 4.5 km/h on level, unimpeded path.

This would suggest that relying solely on the ISRID ring model may lead searchers to plan a larger POA than expected from locally derived data. This is not a critical error as it would be better to slightly overestimate POA rather than underestimate POA and risk drawing a boundary that excluded the missing person. However, deriving ring models based on local data may refine POA techniques and help to better allocate resources. Once data has been collected and formatted, the ring model, is simple to apply using a GIS (or even hand-drawn on a paper map) and can be considered a good starting point for defining a crude search boundary in the first minutes of a search. In all cases, we suggest that ring models should not be used alone without considering other sources of information such as mobility models discussed below. Overall, the ring model still provides an adequate starting point for areas that do not have georeferenced datasets and it is important that the 25% quartiles were similar since these should be the highest priority in the early phases of a search.

Mobility model

The mean T_{min} value for the 130 missing hikers ($T_{min} = 1.5$ h) and the median ($T_{min} = 1.0$) were widely separated by 30 min which may in part be due to the 15 min resolution of the Path Distance Surface. In addition when compared to the ISRID, the Yosemite dataset differed significantly (Table 2; $n = 130$, $\chi^2 = 91.4$, $P < 0.01$). These differences may be due to differences in the way data were collected. In ISRID, mobility values were taken from directly from incident report narratives, whereas Yosemite T_{min} was derived from a modeling process (actual mobility values were not found in incident reports). The ISRID data is inherently biased in that mobility values were estimated and only available from hikers who were found alive (the subjects found deceased on arrival cannot provide this information unless a GPS recorded such data). Analysis of the Yosemite data suggests 50% of hikers should be found in the 1.0 h travel time from the IPP as opposed to the 3.0 h recommendation provided from ISRID. This would suggest that search areas should be prioritized close to the IPP with a maximum distance of approximately 6 h, since hikers seldom are found further than 6 h from IPP (95% < 6 h). Overall, the travel-cost model does present a useful GIS tool for visualizing probability areas and can be created using readily available base data.

Comparison of the ring and mobility model

When the ring model and mobility models were compared statistically, we found a significant correlation between D and T_{min}

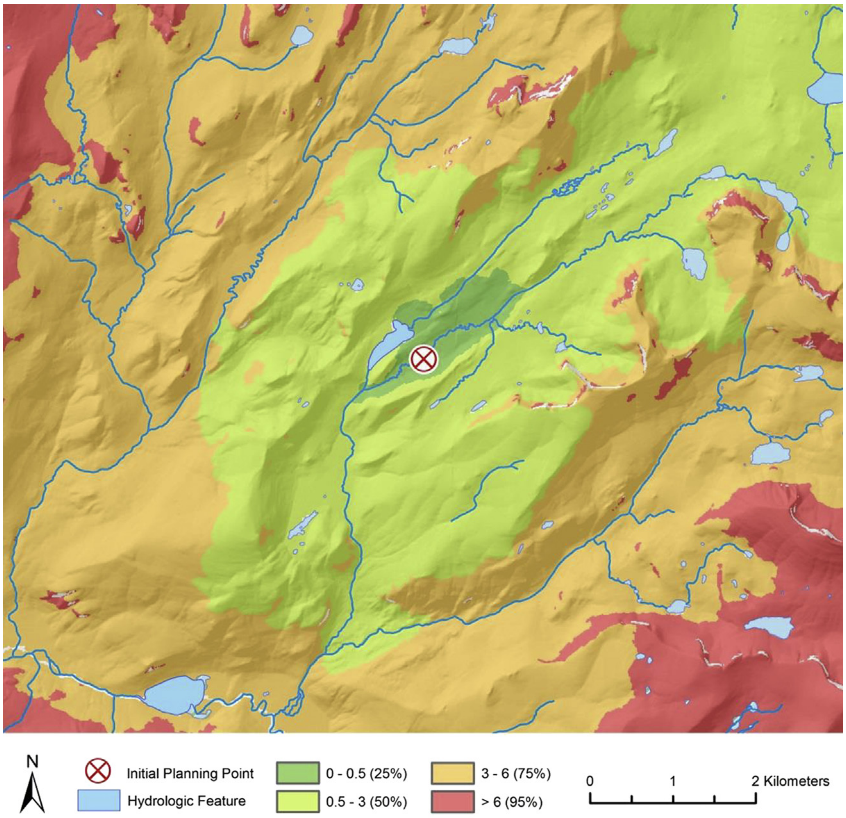


Fig. 8. Travel time given in hours (isochrones) for a hiker based on cost-distance surface derived from geographic factors such as presence and absence of roads, trails lakes, and streams, terrain (slope, cliffs, etc.) and land cover layers. Number of classes (isochrones) relevant to data reported from ISRID mobility table for the hiker category.

(Fig. 9; $n = 130$, $R^2 = 0.96$, $P < 0.01$). This relationship between distance and time is expected, but the scatterplot (Fig. 3) shows that there is a clear limit to the Euclidean distance (D) as a function of T_{min}. If we divide D by T_{min} for each incident, it yields an average value of 2.02 km/h and maximum of 3.75 km/h. This can most likely be explained by two factors: the physical limitations of hiking across varied terrain and the behavior of hikers. From a human geography standpoint, movement is limited by terrain and is represented by the travel-cost model. D/T_{min} will be greatest and approach a maximum where the line between IPP and the location found is across flat terrain and on a trail. Conversely, D/T_{min} will be minimal when the line between IPP and location found is across steep terrain and off-trail. When we consider hiking behavior, D/T_{min} is maximized by a straight line and minimized by a circuitous path that closes in on itself over increasing elevation. We would expect the greatest variation of D/T_{min} in environments with varied terrain and circuitous trails. This is an important geographic phenomenon to understand relative to the use of the ring and mobility models. From a practical standpoint, in mountainous terrain the mobility model will provide greatest contribution of information, whereas in flat terrain with no vegetation the mobility model and ring model would be functionally identical.

The mobility model technique provides a more realistic tool for search boundary delineation because it is based on terrain and human geography. Essentially, if we know the point last seen, time of last sighting, and possible directions of travel, then we can create a potential path area for missing persons using a field-based approach to represent isochrones over the search area. Similar to the mobility models produced in this study, lines of equal travel time (isochrones) are used for studying accessibility in the city of Glasgow (O'Sullivan, Morrison, & Shearer, 2000) and demonstrate a useful tool for visualizing space–time geography. Space–time accessibility is a field of research that is well studied in urban environments by geographers and GIScientists (Kim & Kwan, 2003) and a similar approach has been suggested by other researchers (Lin & Goodrich, 2010; Miller & Bridwell, 2009) for use in WiSAR, yet little research has been done to expand this concept and the current study is the first of its kind to evaluate travel-cost techniques for missing person mobility models using detailed accounts.

It is important to denote that travel-cost techniques do not provide predictive models as they do not take into account the

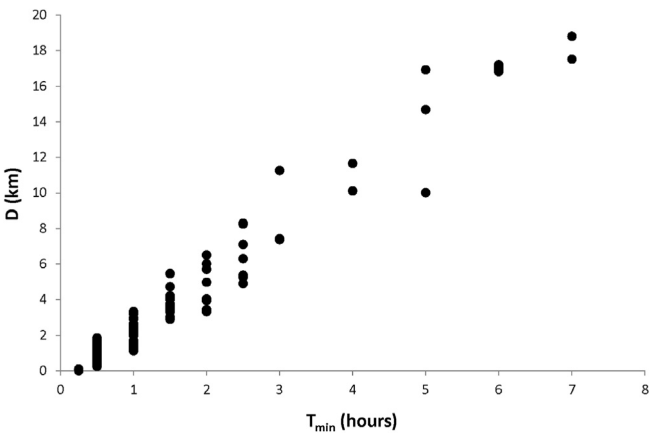


Fig. 9. A scatterplot of isochrones (T_{min}) values at point found versus Euclidean distance (D) between point found and initial planning coordinates. The relationship showed a significant positive correlation ($n = 130$, $R^2 = 0.96$, $P < 0.01$).

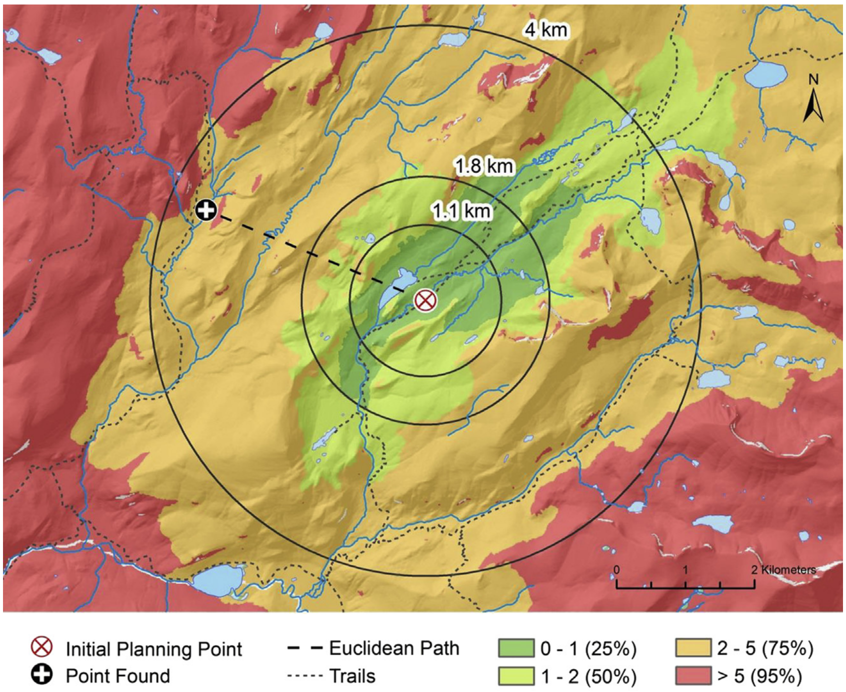


Fig. 10. A map where both the ring and mobility models (isochrones in hours of travel time) derived from local data are incorporated for decision making on probability of area.

behavior of a person (i.e. wandering, returning to IPP, resting) but they do provide some guidance based on information that is known: the landscape the missing person is in and the amount of time they have been missing. Moreover, this model produces isochrones that allow for scenario building if provided in an interactive GIS. For instance, if we have reason to believe a missing person would only be in motion for *x* hours (e.g. due to physical limitations or the arrival of a winter storm that would limit movement), we can indicate suggested probability area as indicated on the map by the isochrones equal to *x* hours. Mobility can be limited by other factors such as incoming weather, darkness, or the behavior of the person themselves (e.g. mental status). As suggested by Pingel (2013), we believe that travel-costs methods can underestimate the cost of travel in hilly and mountainous terrain because these methods do not take into account human perceptions of slope and elevation change (Yang, Dixon, & Proffitt, 1999). However, this is an acceptable bias for developing POA in this use-case as we would rather overestimate than underestimate travel speeds. By using local data as described in this study we can verify the effect of this bias and future research should focus on agent-based techniques to calibrate these models. In the meantime, travel-cost modeling is a tool with visually compelling results that can be used in conjunction with the ring model and other elements of search theory using GISs. We demonstrate this with a historic case where these techniques could have been used to help develop a more well-informed POA if they were available to search planners at the time (Fig. 10).

From this evaluation of the ring model and mobility model, we observed that other datasets can be used in a mountainous area such as Yosemite for forming functional planning areas and delineating the search boundary. One example is watershed boundaries derived from *Federal Standards and Procedures for the National Watershed Boundary Dataset (WBD)* (2012). Preliminary investigation (Doke, 2012) suggests that missing persons rarely are found in watersheds further away than the one in which they were last seen. This fits well within the findings of our research since watersheds are determined by terrain models and are essentially bordered by ridgelines that greatly influence mobility models using travel-cost techniques. Further research should be conducted to determine the usefulness of watersheds in search area planning since it is a readily available GIS dataset.

Conclusions

We have found that developing ring and mobility models from local data will give the best estimate of POA, but that in absence of these data global datasets will provide distance based models that can be used given one understands the constraints in mountainous terrain (a potentially larger POA than would be explained by local data). Overall, there is a great need for more research in the time geography of missing person searches. Efforts should be focused on georeferencing incidents up front in a records management system and then properly analyzing the data on a local level. Then software tools based on GISs can be developed for assisting search managers. While we can learn from a global dataset, local information should be used for actually deriving meaningful probability of area and delineating search boundaries. Once more detailed behavioral data can be obtained, the ring model and mobility model concepts can be combined using more advanced GIS techniques. In the meantime, the authors have begun integrating POA tools into desktop GIS software to learn more about how these concepts apply to actual search incidents. We have identified techniques here that can be used for improving the search for missing persons; however it will require a community of experts in the field of geography working in conjunction with search and rescue personnel to answer the fundamental question: "where is the missing subject?"

References

Adams, A. L., Schmidt, T. A., Newgard, C. D., Federiuk, C. S., Christie, M., Scorvo, S., et al. (2007). Search is a time-critical event: when search and rescue missions may become futile. *Wilderness & Environmental Medicine, 18*(2), 95–101.

Adriaensen, F., Chardon, J. P., De Blust, G., Swinnen, E., Villable, S., Gulinck, H., et al. (2003). The application of "least-cost" modelling as a functional landscape model. *Landscape and Urban Planning, 64*(4), 233–247.

Bailey, R. G., & Ropes, L. (1998). *Ecoregions: The ecosystem geography of the oceans and continents: With 106 illustrations, with 55 in color*. Springer.

Bateman, I. J., Garrod, G. D., Brainard, J. S., & Lovett, A. A. (1996). Measurement issues in the travel cost method: a geographical information systems approach. *Journal of Agricultural Economics, 47*(1–4), 191–205.

Bownds, J. W., Ebersole, M. J., Lovelock, D., O'Connor, D. J., & Toman, R. J. (2007). Win CASIE III: computer aided search information exchange.

Cooper, D. C., Frost, J. R., & Robe, R. Q. (2003). Capability of land SAR procedures with search theory. Retrieved 24.09.13 from http://www.dtic.mil/dtic/tr/fulltext/u2/a511658.pdf.

Doherty, P., Guo, Q., Liu, Y., Wieczorek, J., & Doke, J. (2011). Georeferencing incidents from locality descriptions and its applications: a case study from Yosemite National Park search and rescue. *Transactions in GIS, 15*(6), 775–793.

Doke, J. (2012). *Analysis of search incidents and lost person behavior in Yosemite National Park*. Esri. ArcGIS 10.1.

Federal Standards and Procedures for the National Watershed Boundary Dataset (WBD). (2012). Reston, Virginia.

Ferguson, D. (2008). GIS for wilderness search and rescue. In *2008 ESRI federal user conference*. Retrieved 03.04.12 from http://proceedings.esri.com/library/userconf/feduc08/papers/gis_for_wilderness_search_and_rescue.pdf.

Filipkowska, E., Koester, R. J., Chrustek, R., & Zaród, M. (2012). Lost and found in the Polish Carpathian Mountains. *ArcNews, 34*(3).

Frost, J. R. (1999). Principles of search theory. *Response, 17*(2), 1–23.

Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., et al. (2011). Completion of the 2006 National Land Cover Database for the conterminous United States. *PE&RS, 77*(9), 858–864.

Hill, K. (1999). *Lost person behavior*. Ottawa, Canada: National Search and Rescue Secretariat.

Imhof, E. (1950). *Gelaende und Karte*. Zurich, Switzerland: Rentsch.

Jobe, R. T., & White, P. S. (2009). A new cost-distance model for human accessibility and an evaluation of accessibility bias in permanent vegetation plots in Great Smoky Mountains National Park, USA. *Journal of Vegetation Science, 20*(6), 1099–1109. http://dx.doi.org/10.1111/j.1654-1103.2009.01108.x.

Kim, H.-M., & Kwan, M.-P. (2003). Space–time accessibility measures: a geocomputational algorithm with a focus on the feasible opportunity set and possible activity duration. *Journal of Geographical Systems, 5*(1), 71–91.

Koester, R. (2008). *Lost person behavior*. Search. Charlottesville, VA: dbS Productions.

Koester, R. J., Cooper, D. C., Frost, J. R., & Robe, R. Q. (2004). *Sweep width estimation for ground search and rescue*. Washington.

Koester, R. J., & Stooksbury, D. E. (1995). Behavioral profile of possible Alzheimer's disease patients in Virginia search and rescue incidents. *Wilderness & Environmental Medicine, 6*, 34–43.

Koopman, B. O. (1980). *Search and screening*. New York: Pergamon Press.

Langmuir, Eric (1984). *Mountaincraft and leadership. Official handbook of the Mountain Leader Training Boards of Great Britain and Northern Ireland*. Edinburgh, Scotland: Britain & Scottish Sports Council.

Lin, L., & Goodrich, M. A. (2010). A Bayesian approach to modeling lost person behaviors based on terrain features in wilderness search and rescue. *Computational and Mathematical Organization Theory, 16*(3), 300–323.

Magyari-Sáska, Zs., & Dombay, Ş. (2012). *Determining minimum hiking time using DEM*. Geographia Napocensis, anul VI. Nr. 2, Cluj-Napoca.

Mattson, R. J. (1976). Establishing search areas. *Search and Rescue Magazine, Spring 1976*, 7–8.

Miller, H. J., & Bridwell, S. A. (2009). A field-based theory for time geography. *Annals of the Association of American Geographers, 99*(1), 49–75.

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 (Bethesda, Md.: 1985), 93*(3), 1039–1046. http://dx.doi.org/10.1152/japplphysiol.01177.2001.

Nikolakaki, P. (2004). A GIS site-selection process for habitat creation: estimating connectivity of habitat patches. *Landscape and Urban Planning, 68*(1), 77–94.

NSARC. (2011). Land search and rescue addendum to the national search and rescue supplement to the international aeronautical and maritime search and rescue manual. version 1.0 http://www.uscg.mil/nsarc.

O'Sullivan, D., Morrison, A., & Shearer, J. (2000). Using desktop GIS for the investigation of accessibility by public transport: an isochrone approach. *International Journal of Geographical Information Science, 14*(1), 85–104.

Pingel, T. J. (2013). Cartography and geographic information science. *Cartography and Geographic Information Science, 37*(2), 37–41.

Ralston, H. J. (1958). Energy-speed relation and optimal speed during level walking. *Internationale Zeitschrift für Angewandte Physiologie Einschließlich Arbeitsphysiologie, 17*(4), 277–283. http://dx.doi.org/10.1007/BF00698754.

Sherrill, K. R., Frakes, B., & Schupbach, S. (2010). *Travel time cost surface model: Standard operating procedure*. Natural resource report NPS/NRPC/IMD/NRR-2010/238. Fort Collins, Colorado.

Stone, L. D. (1989). What's happened in search theory since the 1975 Lanchester Prize? *Operations Research, 37*(3), 501.

Stone, L. D. (2007). *Theory of optimal search* (2nd ed.). In *Mathematics in science and engineering*. New York: Academic Press.

Syrotuck, W. G. (2000). *Analysis of lost person behavior*. Mechanicsburg, PA: Barkleigh Productions, Inc.

Theodore, J. (March 2009). When every second counts. *ArcNews*, 66–69.

Tobler, W. (1965). Non-isotropic geographic modeling. *Information systems*.

Tobler, W. (1991). Non-isotropic geographic modeling. In W. Tobler (Ed.), *Geographic information systems in the social sciences*. Santa Barbara.

Van den Broek, M., Brederode, E., Ramírez, A., Kramers, L., Van der Kuip, M., Wildenborg, T., et al. (2010). Designing a cost-effective CO_2 storage infrastructure using a GIS based linear optimization energy model. *Environmental Modelling & Software, 25*(12), 1754–1768.

Whitley, T., & Hicks, L. (2003). A geographic information systems approach to understanding potential prehistoric and historic travel corridors. *Southeastern Archaeology* (those 1994).

Wieczorek, J., Guo, Q., & Hijmans, R. (2004). The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. *International Journal of Geographical Information Science, 18*(8), 745–767.

Winter, S., & Yin, Z.-C. (2010). Directed movements in probabilistic time geography. *International Journal of Geographical Information Science, 24*(9), 1349–1365.

Yang, T. L., Dixon, M. W., & Proffitt, D. R. (1999). Seeing big things: overestimation of heights is greater for real objects than for objects in pictures. *Perception, 28*(4), 445–467.