Terrain Based Probability Models for SAR

Matt Jacobs
matt@mattj.net

May 15, 2015

Contents

1 Abstract

In order to examine the effect of terrain on search and rescue (SAR) find locations, incidents from the International Search and Rescue Incident Database (ISRID) were compared against a variety of geospatial data. Unlike most SAR research to date, find locations were examined independent of a subject's last known location or possible travel route. Notable variations in find probability relative to terrain were observed, as well as differences in the find locations of injured and uninjured subjects.

2 Introduction

Much of our understanding of land search theory derives from work done by the US Navy and Coast Guard, originating with submarine hunting during WWII and subsequently applying that knowledge to maritime SAR. While the techniques developed by these agencies are generally applicable to land SAR, one fundamental difference is the presence of terrain and manmade features. Open water searches have a smooth probability density distribution, in which small changes in position result in small changes to the probability of a find. In land SAR, by contrast, the chance of a find can vary significantly with small changes in position.

Broadly speaking, land searches are accomplished using two techniques: linear feature searches, such as those along a trail or stream, and area searches, in which searchers walk a uniformly spaced grid. On land, searcher speed and spacing will also vary throughout an operation, with fast searching for a responsive subject generally giving way to slow, careful searching for an unresponsive one as time goes on.

Due to a lack of linear features, open water searches are primarily grid based. A narrower variety of techniques are generally employed on each search, as a responsive individual in open water is only moderately easier to spot from the air than an unresponsive one. Additionally, open water conditions are more consistent - swell height does not change over short distances to the degree that vegetation does.

In the presence of smooth distributions for both subject location and search conditions, the most effective method is a uniformly spaced gridded area search. To maximize the efficiency of this technique, the maritime SAR community has expended significant effort developing smooth large-scale probability distributions based on factors such as wind and current. There has also been considerable research on the distances over which various objects can be seen for given ambient conditions, resource types and searcher speeds.

Mirroring this, there has been effort within the land SAR community to develop smooth large-scale probability distributions based on factors such as expected travel distance, elevation change and dispersion angle. There has also been some research, primarily funded by the Department of Homeland Security (DHS), on the distances at which various objects can be seen in different types of vegetation. If for no other reason than funding, neither of these subjects have been as thoroughly researched as in maritime SAR.

Even so, despite the major differences between land and maritime environments, comparatively less

effort has gone into researching the impact of terrain on subject behavior. I am not aware of any research that quantifies those effects; Robert Koester's *Lost Person Behavior* lists find percentages for various features and track offsets, but without knowing the composition of the search areas reported on, those numbers can not be translated into find probabilities.

I addressed this by looking at the relationship between find locations and a variety of manmade and natural features. Where most research to date has focused on the relationship between the find location and a person's last known point (LKP), I chose to examine find locations in isolation. This decision was partially driven by the limited number and quality of reported LKPs, but also because questions such as "how likely is someone to be found in a drainage" can be answered independent of a subject's possible travel route.

For each subject category presented in *Lost Person Behavior*, a statistic such as distance travelled is listed if it has been reported for at least 14 incidents. While this may be adequate for reporting on population statistics, the methods used in this paper require more data. For example, determining a stream's probability from the number of near-stream finds is analogous to determining how a die is weighted by rolling it repeatedly. Additionally, I wanted enough incidents to analyze not only the dataset as a whole, but also to subdivide on factors such as subject status and distance travelled.

Because most categories lacked a sufficient number of find locations, only the largest group of categories - hikers, hunters and gatherers - was used. Additionally, the small proportion of urban finds were discarded, and only backcountry incidents were analyzed.

Terrain features examined include roads, trails, streams, lake shores, coastlines, elevation, slope angle, land cover, ridges and drainages. Non-linear manmade features such as buildings and trailheads were not considered. It seems plausible that features might follow some kind of tiered hierarchy, with for example streams only being relevant to off-trail finds. However, as there is no evidence to support this, the dataset was generally not filtered when looking at features presumed to be lower probability.

2.1 Limitations

Three quarters of the find locations studied came from Oregon, with the remainder from New York and Arizona. Additionally, most analysis was focused on backcountry incidents involving the ISRID hiker, hunter and gatherer groups. The applicability of these results to other locations, terrain types and subject categories is an open question.

While this paper falls under an area of research often referred to as lost person behavior, that name is a bit misleading. No attempt has been made to differentiate between subjects who injure themselves in a drainage and stay put, and those who wander downhill into a drainage once injured.

This paper explores a relatively new avenue of research; the results presented here are ripe for further analysis using improved techniques and larger sample sizes. While I hope the rough sketch will withstand the test of time, it seems inevitable that some of the finer-grained conclusions may eventually be invalidated.

2.2 Search Theory

The classic approach to search management is to:

1. Establish the search area
2. Segment the search area
3. Assign probabilities to each segment

Every search begins with an Initial Planning Point (IPP) representing the subject's last known position. The search area is established as a circle centered on this point, with a radius derived from several sources including historical lost person behavior data from ISRID. This search area is then divided into non-overlapping searchable areas called segments.

Each segment is assigned a Probability of Area (POA) reflecting the likelihood that the subject is within that area. A parallel concept to POA is probability density, or PDEN. The units for PDEN are (probability) / (unit area), and it is generally expressed in real units by dividing a segment's POA by its size, e.g. percentage probability per square kilometer.

Actual search efforts are characterized by a conditional Probability of Detection (POD), the probability that a team would have found the subject if the subject were actually within the searched segment. In uniform terrain, POD can be determined using the range at which searchers are likely to see the subject (effective sweep width) and distance travelled by searchers.

Underlying these concepts is the assumption that within each segment, all points are equally likely to contain the subject - mathematically, segments are assumed to have uniform PDEN. In theory, features that cause large steps in PDEN (e.g. trails) have already been searched during the initial (hasty) phase, and gradual PDEN changes are handled through well-placed segment boundaries.

However, there is no established standard for identifying high-probability hasty features. The two major search management textbooks, NASAR (p 204) and ERI (p 253), provide lists that collectively include roads, trails, streams, rivers, creeks, drainages, ridges, lines of little resistance, power lines and clearings, as well as unspecified attractions, hazards and likely spots. While this list far outstrips the resources available in a typical hasty search, no indication is given as to their relative priority.

As a search progresses, some of these hasty features are broken out into linear segments for high POD searching, but there is no established guide for weighting their POA as compared to nearby areas. Further, there is no evidence-based standard from which to develop area segments that will have a uniform PDEN distribution.

Since this paper is looking at find locations without regard to an established search area, it is impossible to express PDEN in real-world units. Instead, PDEN is presented as (% probability) / (% search area). As any randomly chosen 1% of the search area has a 1% chance of containing the subject, all terrain has a default PDEN of 1. In this sense PDEN can be used a multiplier - a PDEN of 4 for areas within 100' of lakes would mean they are 4x more likely to contain the subject than randomly chosen terrain.

The actual POA for a given lake shore can only be determined by combining the terrain based models presented here with traditional distance based behavior models.

3 Techniques

Find locations were compared against two classes of data: raster (a grid of values, like an image) and vector (points, lines and polygons). Raster data includes elevation and land cover; vector data includes roads, trails and streams.

For both data types, find locations need to be considered in the context of their surrounding terrain; a finding that half of all subjects are located within 100m of a road is more useful in a wilderness area than one crisscrossed with logging roads. For an incident database containing many find locations within a constrained geographic area, such as a small national park, it might be possible to compare find locations against the park as a whole. For example, comparing the percentage of hikers found in clearings against the percentage of parkland occupied by clearings would help determine how strongly clearings predict the find location. Although he was not looking at the predictive values of terrain, see Jared Doke's 2012 paper for an example of such a constrained-area analysis.

The ISRID dataset is too geographically sparse for this approach, and the results would be biased towards the terrain in which people go missing. In short, knowing that most hikers are found in the forest would help you search the state for a missing hiker's car, but is of little predictive value once an IPP is established.

An alternative approach is to generate a uniformly distributed set of points ("sample points") within a given radius ("surrounding circle") of each find location, and then compare the find location against the sample points. In this paper, the median IPP-find distance (2km) was used as the surrounding circle radius. Finds can be compared directly against their individual surrounding circles, or against all surrounding circles as a whole.

Using the latter approach, PDEN for a given criteria (e.g. "near trails") can be determined using (% of find locations matching criteria) / (% of sample locations matching criteria). Where road and trail datasets are incomplete, the derived PDEN will still be valid as long as there is no correlation between find locations and dataset errors. As an extreme example, if half of all roads are missing from the dataset, half of the on-road finds would be misclassified as off-road, but because a similar percentage of search terrain would likewise be classified as off-road, the ratio would remain the same.

3.1 Linear Features

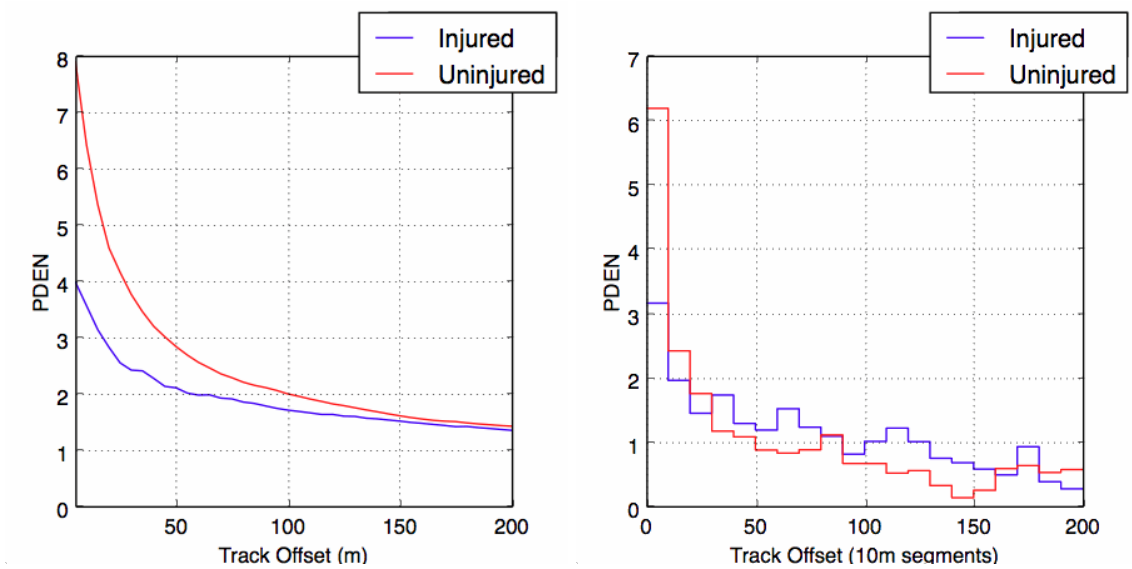
One of my primary goals was to answer questions like "what percentage of people are found on trails?" and "what is a trail's PDEN?". Implicit in both of these questions is the ability to look at a point and categorize it as being either "trail" or "not trail". This is a challenging question in itself, but further complicated by positional errors in both reported find locations and the datasets those locations are compared against. While some finds 10 meters from a road may be on-road finds with positional errors, others may be someone who suffered a stroke, walked off the road and is not easily detected via hasty search.

Existing SAR research has approached this problem using the concept of track offset. Simply put, a track offset is the shortest distance from a point to a linear feature. It can be used to assign a number to a single point ("this location is 20m from the nearest road"), or to describe an entire set

of points ("all points within 20m of a road"). Searching a feature to a 100m track offset actually requires searching a 200m wide strip - 100m to the left of the feature, and another 100m to the right. While *Lost Person Behavior* only applies track offset to finds not actually on a particular feature, points can never truly be "on" the 1-dimensional line data used in this study. Instead, all points are given a track offset, including those incredibly close to linear features.

Back to the question of trail PDEN and "on trail" finds, the problem can be visualized by plotting PDEN against track offset. A cumulative PDEN plot (Figure 1a) shows PDEN for the entire track offset; a cumulative PDEN of 2 at a 50m offset would mean that if you search everything within 50 meters of a feature, you can expect a PDEN of 2. The cumulative graphs presented here start at a track offset of 5 meters and have additional data points at 5 meter intervals, out to 200m.

A segmented PDEN plot (Figure 1b) shows PDEN for incremental changes in track offset; assuming 10 meter increments, a segmented PDEN of 2 at a 50m offset means that if you search only between 40 and 50 meters of a feature, you can expect a PDEN of 2. The segmented plots presented here generally use 10 meter increments, but may use larger ones for small datasets.



(a) Cumulative *(b) Segmented*

Figure 1: Example PDEN plotted against track offset from manmade linear features, by subject status.

While segmented plots provide a better picture of the way PDEN changes with track offset, they are inherently noisier than cumulative ones. It's important to look a the overall trend and not the small fluctuations from one incremental track offset to the next. This noise can also make it harder to compare multiple lines - for example, injured and uninjured subjects - on the same plot. Ultimately both segmented and cumulative plots are useful in forming a coherent picture of the data.

It can also be useful to look at how PDEN changes in relation to a second variable such as distance from the IPP. This can be accomplished by picking a fixed track offset, for example 40 meters, and plotting PDEN within that offset against the second variable. Because PDEN can only be calculated from a collection of find locations, each point on the X axis represents a range of values

(a "window"). A PDEN of 2 for a distance of 1000m would mean for all finds within a distance of 1000 ± the window size, the 40m cumulative PDEN was 2. To reduce noise, the windows overlap by 50%; if the X axis has 500m increments, the window size is 750m.

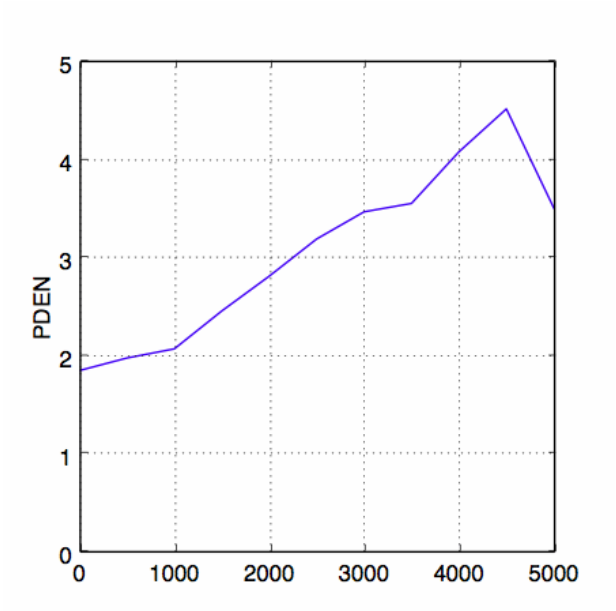


Figure 2: Example PDEN for a 40m track offset as a function of distance from the IPP.

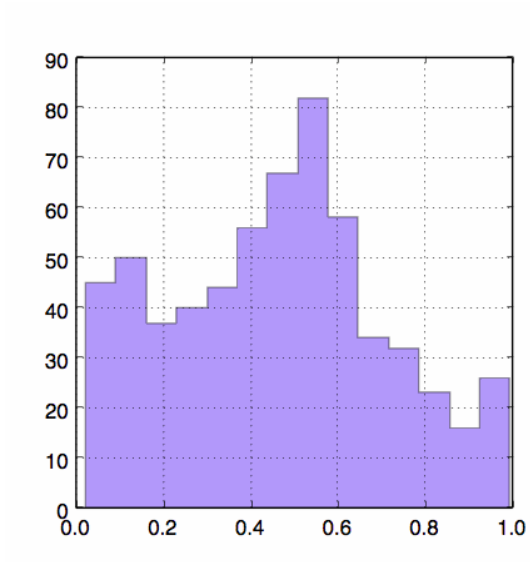
3.2 Raster Data

Track offset is not a meaningful concept for raster data like elevation and land cover. While it is possible to determine each point's distance from a nonlinear feature such as a summit or meadow, this paper instead uses the direct value a point lies on, and not its offset from some other feature.

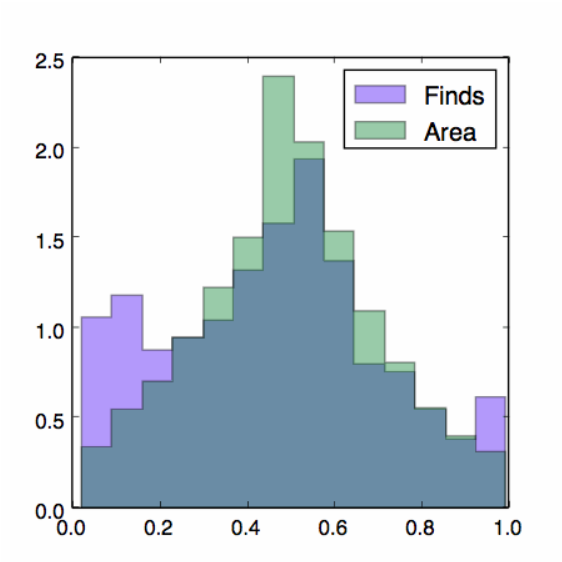
While a point's raw numerical value can be used in some cases, it's often useful to look at that point in relation to its surrounding terrain. Elevation is a good example, as it's meaningless to compare raw find elevations between coastal and mountainous areas. One alternative is to rank points against their surrounding terrain on a percentile basis; a 0.05 (5%) rank would mean that a point is lower than 95% of its surrounding circle.

If find locations and sample points are all given percentiles basis values, PDENs can be generated. For example, if 20% of find locations have a percentile basis below 0.1, but only 5% of search terrain does, then those low-lying areas have a PDEN of 4 (20% / 5%). As with track offset, percentile basis PDEN can be shown on a segmented plot. An increment of 0.1 (typical for this paper) would show PDEN for percentile basis values of 0-0.1, then 0.1-0.2, etc.

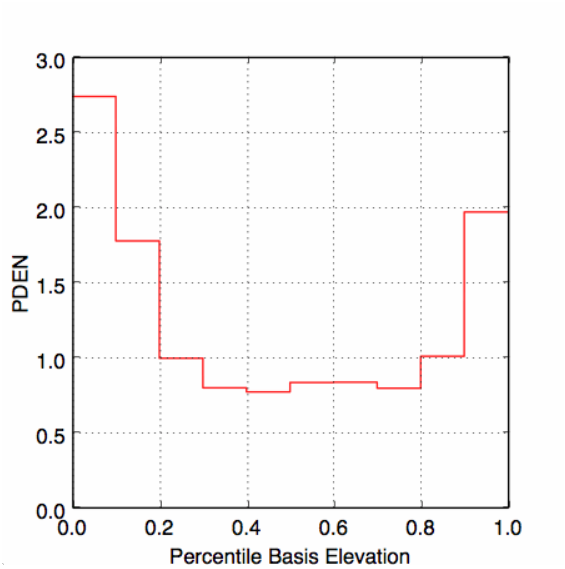
Figure 3 illustrates these concepts using percentile basis elevation. More finds are located mid-slope than at high or low points, suggesting that mid-slope locations might be better places to search. However, once find locations are compared to nearby sample points, it becomes obvious that mid-slope points in general outnumber high and low points, and that the extremes of elevation are actually better places to look.



(a) Histogram of percentile basis elevation for find locations.



(b) Percentile basis elevation for find locations (blue) compared to sample points (green).



(c) Segmented PDEN for percentile basis elevation

Figure 3: Figure 3a shows that percentile basis elevations of 0.4-0.6 are most common for find locations. Without context, this could lead one to conclude that midpoints have a higher PDEN than high and low points. Figure 3b superimposes this graph against all search terrain, showing that find locations are more likely to be high or low points than randomly chosen terrain. Figure 3c plots PDEN against percentile basis elevation, showing the opposite of what one might conclude from Figure 3a.

3.2.1 Significance

In any observational study or experiment, there is always a possibility that seemingly interesting results could have happened purely by chance. Typically a test statistic, such as a normal curve, is used to determine how likely this is to have happened. If it is unlikely that the results could have happened by chance, they are considered statistically significant; the cutoff for significance is generally 5%, i.e. a 1/20 chance that the results are random noise.

In this paper, sample points are used as a control. If 10% of search terrain is near streams, and streams have no effect, then one would expect roughly 10% of finds to be near streams. Due to random chance, the number of near-stream finds is unlikely to be exactly 10%, especially for smaller sample sizes. Instead, the distribution of find percentages within a fixed track offset of streams will follow a binomial distribution, with each incident representing a separate binary experiment.

The number of finds matching a given criteria (e.g. near streams) is only significant if it exceeds the relevant binomial distribution's 95th percentile. If 10% of the search terrain is near streams and 100 find locations are examined, this distribution will have number of tries n = 100 and probability p = 0.1; in this case, the 95th percentile works out to 16. Put another way, if you made a game out of randomly throwing 100 darts at a map, 19 out of 20 times you would hit streams with fewer than 16 darts.

Another useful way to measure results is margin of error. Margin of error takes an outcome (percent of finds near streams, percent of voters in favor of a candidate) and provides the range you would expect 19 out of 20 repeated polls or experiments to lie within. Keeping with the example above, if 16 out of 100 finds are located near streams (n = 100, p = 0.16), the margin of error is roughly 7%, and the 95% confidence interval runs between 9% and 23%.

Unlike most regression analyses, the significance test described above is one-tailed (i.e. a 5% chance of being over the significance level). The margin of error, however, is two-tailed (2.5% chance of being under the margin and 2.5% chance of being over the margin). Even though the lower bound on the margin of error is below the 10% of terrain occupied by streams, the results are still be significant.

In keeping with standard practices, probability densities will be reported ± the standard error, which is 1/2 of the margin of error. A PDEN of 5±1 means that if it were somehow possible to roll back time and repeat the studied searches hundreds of times, then 2/3 of the time you would expect PDEN to fall between 4 and 6.

A final question examined in this paper is the relationship between two groups: if a greater percentage of group A is found near streams than group B, are the results significant or just noise? Rather than being determined by overlapping margins of error, this is instead evaluated by checking if the difference between the means of each group is greater than $1.96\sqrt{{SE_A}^2 + {SE_B}^2}$, where SE is the standard error for each group.

In this paper, PDEN margin of error was determined by establishing the margin of error for find percentages and then dividing by the percentage of search terrain. The percentage of search terrain matching a given criteria is assumed to be accurate and not assigned its own confidence interval.

4 Results

4.1 Incident Data

This paper looks at incidents from the International Search and Rescue Incident Database (ISRID) that lie within the contiguous US (CONUS); only incidents with coordinates for the find location are used. Some of those incidents also have coordinates for an initial planning point (IPP). While the IPP is nominally a subject's last known location, there is much variation in the quality of IPPs reported to ISRID. An IPP may be the point at which someone wandered off from a group, a trailhead, or even a residence. The variable IPP quality, combined with the low number of reported IPPs, led me to focus primarily on find locations and not on paired IPP and find coordinates.

The CONUS dataset has roughly 2200 incidents with reported find locations. Most incidents are from Oregon, with several hundred each from New York and Arizona. Of those, 1636 are on-land cases with an incident type of search (subject location not known at beginning of incident) or rescue (subject location known at beginning of incident); the remainder are mostly on-water incidents, with some aircraft crashes and other categories. Based on land cover data (see section 4.6), 181 incidents were discarded due to having surrounding circles that were at least 25% developed. The 1455 (89%) remaining backcountry incidents were retained for further evaluation.

Each incident has an ISRID-provided status, also called outcome, of either well, injured or DOA. Due to limited numbers of injured and DOA subjects, they were examined together and are simply referred to as "injured" within this paper. While injured spans all major injuries, from walking wounded to fully unresponsive, status is assumed to have some correlation to mobility and responsiveness, as neither of those are tracked within ISRID.

ISRID also provides an ecoregion domain for each incident (polar, temperate, dry or tropical), and *Lost Person Behavior* relies heavily on ecoregion for modeling subject behavior. Because most of Oregon and all of New York are temperate, and all of Arizona is dry, ecoregion is heavily tied to US state within the examined dataset. To prevent over-generalization, results will generally be broken down by state rather than ecoregion.

4.1.1 Backcountry Incidents

Of the 1455 backcountry incidents examined, 903 (62%) were searches and 552 (38%) were rescues, i.e. the subject location was known at the beginning of the incident. Searches had significantly lower injury rates (81% well, 9% injured, 10% DOA) than rescues (44% well, 51% injured, 5% DOA).

Each incident has an ISRID-provided subject category, such as hiker or hunter. To keep sample sizes large, similar categories were evaluated together as category groups. At 622 incidents, the largest backcountry category group is hikers, hunters, gatherers and runners, who roughly represent able-minded adults traveling on foot in non-technical terrain; they are collectively referred to as the foot category group. The next largest category group is vehicles, with 297 backcountry incidents, followed by skiers and snowboaders (79), children (55) and despondents (47). All remaining category groups comprise 355 backcountry incidents.

As Figure 4 shows, PDEN for roads and trails varied by category group, with vehicles having the highest PDEN and despondent individuals having the lowest. It also varied by subject status (Figure 5), with uninjured subjects having a higher road and trail PDEN than injured ones.

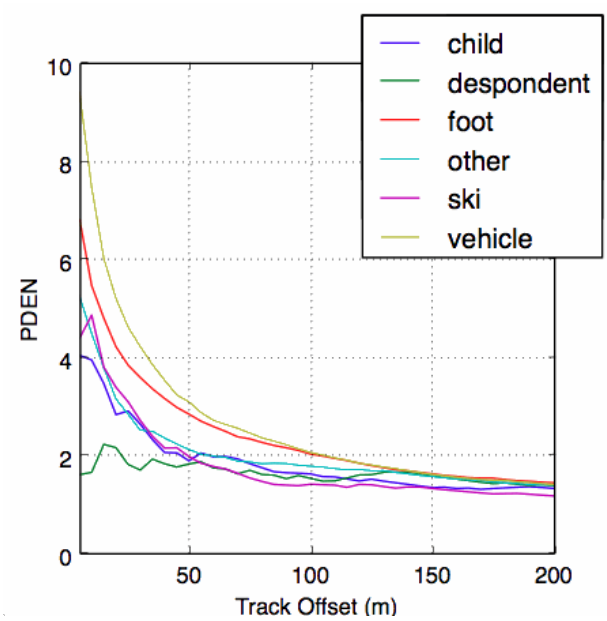


Figure 4: Cumulative road and trail PDEN, by subject category, backcountry incidents

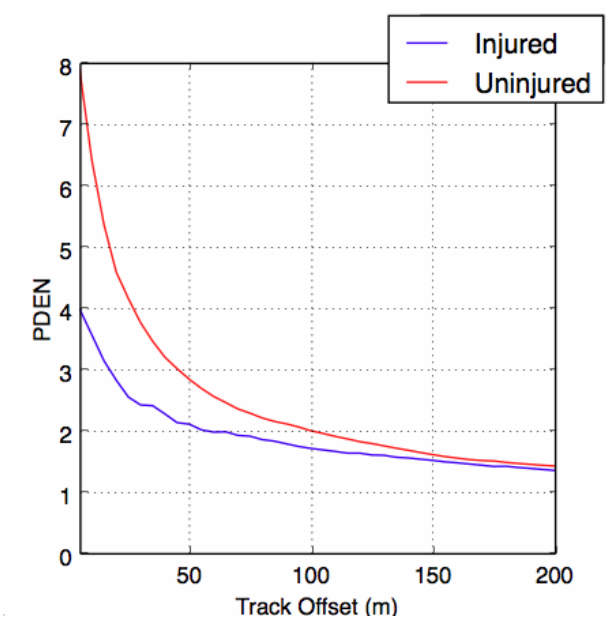


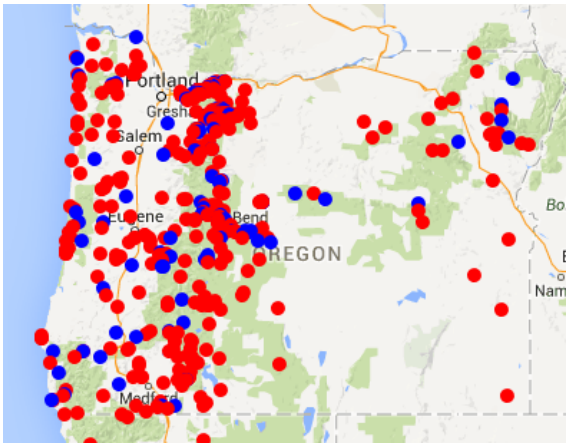
Figure 5: Cumulative road and trail PDEN, by subject status, backcountry incidents

Throughout this paper, only the foot category group was examined in detail.

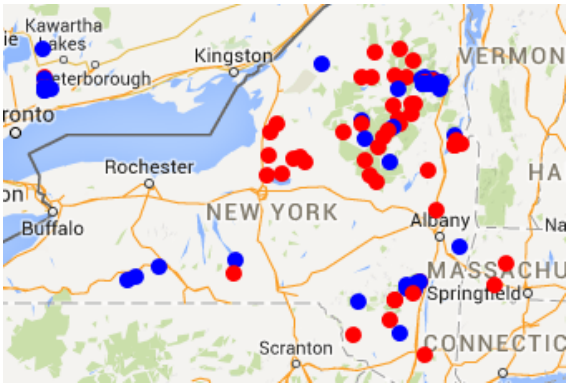
4.1.2 Foot Incidents

Research was focused on backcountry incidents from the foot category group, comprising 622 find locations. Incident types were 71% searches (subject location not known at beginning of incident) and 29% rescues (subject location known at beginning of incident). Subject categories within the group were 74% hikers, 14% hunters, 11% gatherers and 1% runners.

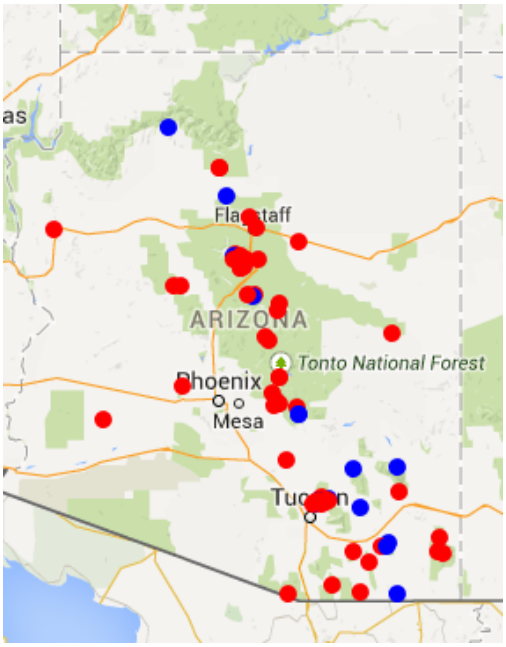
For reference, the locations of these incidents are shown in Figure 6. Incident locations are color coded by status, with uninjured subjects represented by red dots and injured or deceased subjects represented by blue dots.



(a) Oregon



(b) New York



(c) Arizona

Figure 6: Backcountry foot incident locations, color coded by status. Uninjured subjects are red, injured ones are blue.

4.2 Roads and Trails

Roads and trails, collectively referred to as manmade linear features, were sourced from two datasets. One is OpenStreetMap (OSM), an open source, publicly editable map database. Features in it have been both drawn by hand and added in bulk from government data. As a result, OSM data comprehensiveness varies geographically; some areas contain every logging spur, and others have almost no forest roads.

Within national forest lands, OSM data was supplemented by the US Forest Service's FSTopo Transportation dataset. Because the two datasets were used alongside each other, and because mapping errors cause the same feature to appear in slightly different locations in each dataset, some areas appeared to be more road and trail dense than they actually are. At the same time, many features are missing from both datasets entirely, especially newer trails or public land managed by agencies other than the Forest Service.

For both roads and trails, mapping errors mean that real-world track offsets are probably lower than reported here, especially for close by finds. It is also likely that trails are less accurately mapped than roads, with the difference between real-world and computed trail offsets being comparatively greater.

4.2.1 Roads

Across both searches and rescues, road PDEN was found to vary by subject status, with uninjured subjects more likely to be found on or near a road (Figure 7). PDEN decreased more rapidly with track offset for uninjured subjects than injured ones, suggesting that injured subjects are comparatively more likely to be found near, rather than on, roads.

Uninjured subjects had a road PDEN of 4.5±0.4 at a 20m track offset and 3±0.2 within 40m. Injured subjects had a small but still significant ($p = .02$) PDEN of 1.5±0.3 at a 40m offset.

4.2.2 Trails

Injured subjects appeared to have a flatter distribution, with proportionately more finds at track offsets of 20m-40m. Although the 20m-40m PDEN difference was statistically significant, it is also a cherry-picked result and could easily be due to chance. Although trail PDEN is shown here for all incidents, the pattern was similar when examining only off-road (> 40m track offset) find locations.

Uninjured subjects had a trail PDEN of 8±1 at a 20m cumulative offset and 5.5±0.5 at 40m. Injured subjects had a PDEN of 7.5±1 at 40m.

4.2.3 Other Factors

Manmade linear feature PDEN generally increased with both vertical relief and IPP-find distance (Figure 9). As an example, near-manmade finds increased from 30% of incidents within 2km of the IPP to 53% beyond 3km, and from 32% in areas with less than 1000' of relief to 46% elsewhere.

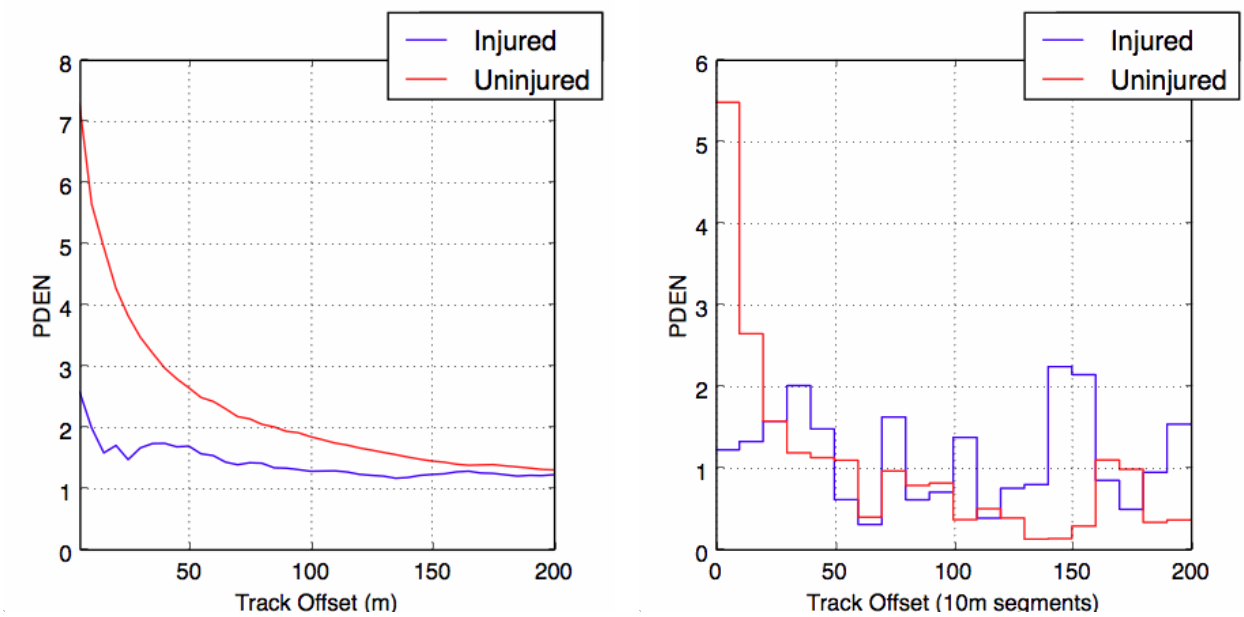


Figure 7: Road PDEN. Uninjured subjects had a high road PDEN that dropped off more rapidly with track offset than injured ones.

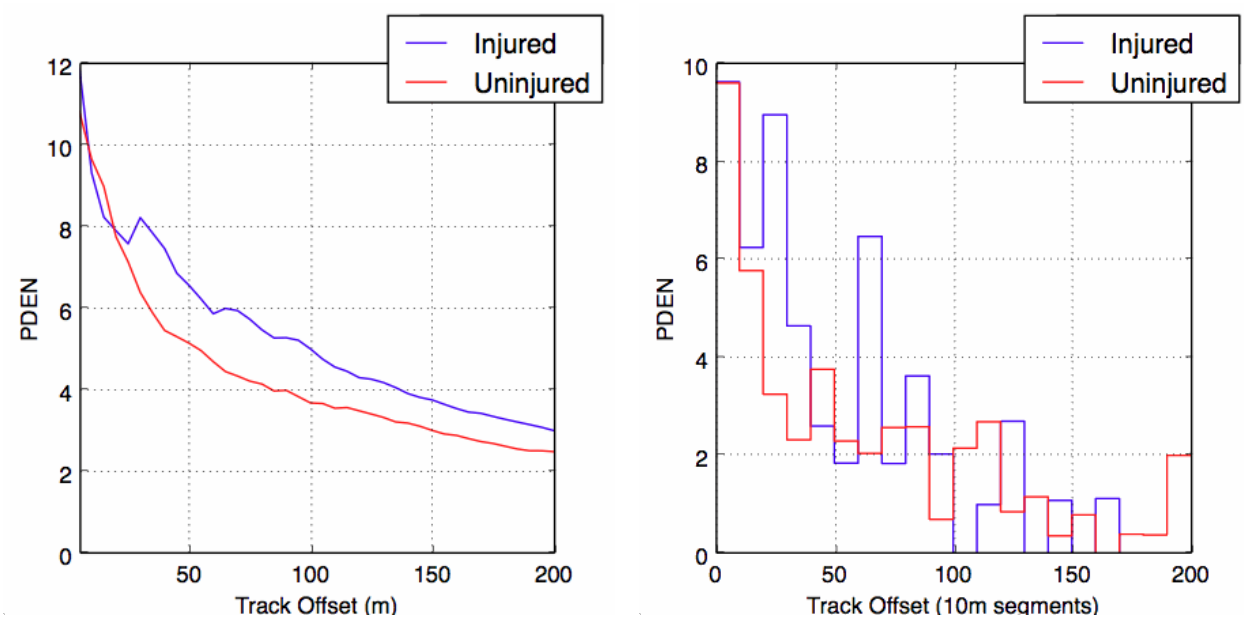


Figure 8: Trail PDEN. Probability was not affected by subject status to the same degree as road PDEN.

Although these trends ceased at the extremes of both relief and distance, sample sizes at those extremes were small.

In terms of find probability, the IPP was of no predictive value beyond its use in standard distance-

based models. The IPP's proximity to manmade features had little correlation to IPP-find distance (Figure 23b) or the find location's proximity to manmade features (Figure 11). The increase in manmade PDEN with distance also occurred for both on and off trail IPPs.

The influence of subject status on PDEN at wider track offsets (i.e. injured subjects being near features rather than on them) became more apparent when discarding rescues and looking only at searches (Figure 10a). However, it remains unclear whether this is an appropriate technique; often the only difference between a search and a rescue is the subject's ability to summon help.

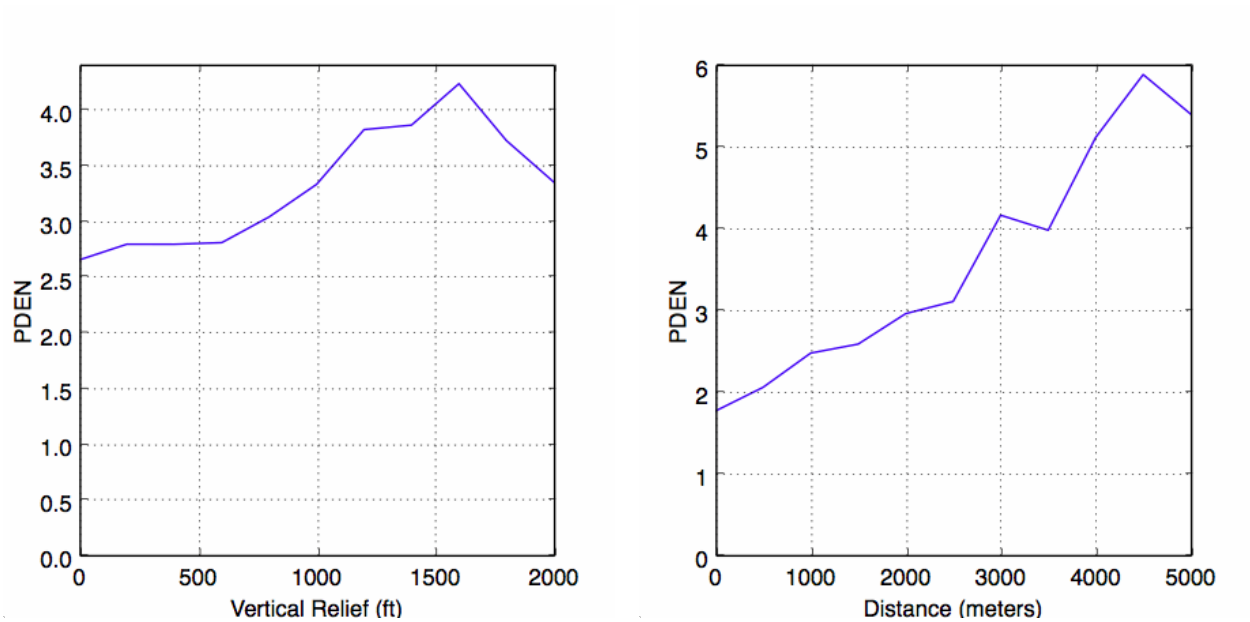


Figure 9: Effect of other factors on manmade linear feature PDEN. PDEN generally increased with both vertical relief (left) and IPP-find distance (right), although it declined at the extremes of both.

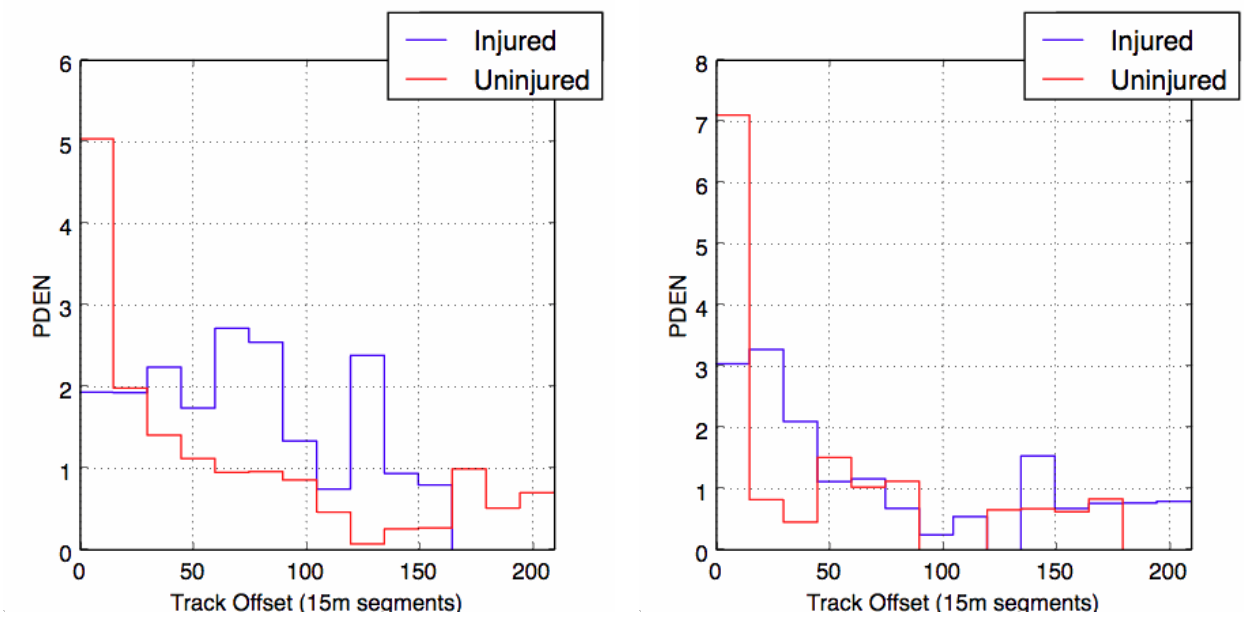
4.2.4 By State

Looking at individual states, roads had noticeably lower PDEN in New York than in Arizona and Oregon. New York also had no off-road incidents within 40m of a mapped trail. This is likely due to data coverage issues, especially since the Adirondacks are managed at the state level and not mapped by the Forest Service, but the discrepancy may be worth further investigation.

4.2.5 Summary

Trails had a significantly higher PDENs than roads, especially for injured subjects. Since only foot incidents were considered, this seems reasonable - people are more likely to hike on trails, and drive when roads are available. However, I would not use this finding as a reason to prioritize trails over roads when they are both likely travel paths.

It is probable that injured subjects were comparatively more likely to be near - rather than on - roads and trails, but this finding warrants further investigation.



(a) Searches *(b) Rescues*

Figure 10: Manmade linear feature PDEN broken down by incident type. Searches had lower PDEN than rescues. Searches also appeared to have a flatter injured distribution than rescues, although this may be an artifact of sample size. Only 26 (out of 47) injured search locations were within 100m of a manmade feature.

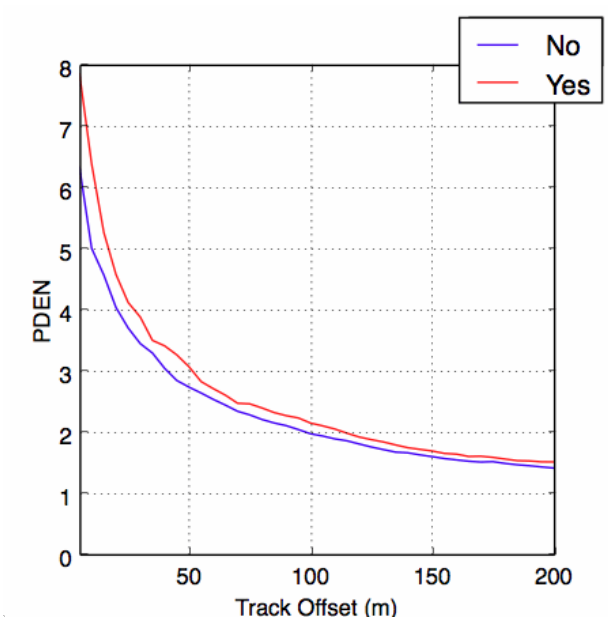
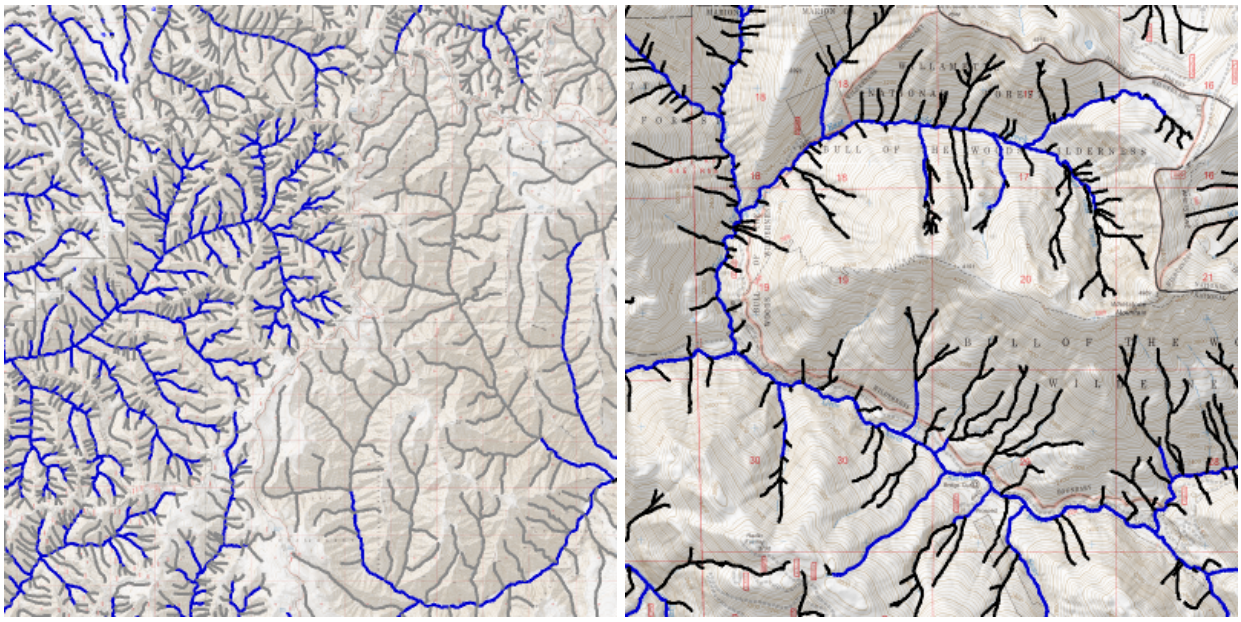


Figure 11: Manmade linear feature PDEN grouped by whether the IPP was near a road or trail. The IPP's proximity to manmade features had little impact on the find location's proximity

4.3 Water

The National Hydrography Dataset (NHD) maps surface water within the contiguous US. Features studied included both polygonal water bodies (lakes and ponds) and linear flowlines (streams, rivers and canals). To maintain reasonable sample sizes, coastlines and all water body boundaries, regardless of size, were counted as lake shores.

The NHD classifies flowlines as either perennial, intermittent or ephemeral. While these may be reliable indicators of a stream's water flow, they do not reflect its terrain prominence. In some locations intermittent and ephemeral streams are limited to shallow drainages, while in others they comprise deep, prominent canyons (Figure 12a).



(a) By NHD classification: perennial (blue) and intermittent or ephemeral (gray).

(b) By Strahler Number: at least 5 (blue) and less than 5 (black)

Figure 12: Color-coded example NHD flowlines.

To provide a greater level of consistency relative to terrain, NHD flowlines were assigned a Strahler number (http://en.wikipedia.org/wiki/Strahler_number) based on the number of upstream junctions. Flowlines with a Strahler number of at least 5 were counted as streams, with the remainder called capillaries. As Figure 12b illustrates, this is more consistent than the NHD classification system but still imperfect.

4.3.1 Lakes

Only 25 incidents were within 40m of a lake or coastline, giving those features a PDEN of 2.2±0.4. However, because a significant portion of the surrounding circle for any near-lake find will also be on or near that lake, this number may be artificially low. There were not enough incidents to further subdivide by category, subject status or other factors.

4.3.2 Streams

Injured subjects had a significantly higher stream PDEN than uninjured ones (Fig 13). Uninjured stream PDEN was 2.2±0.3 at a 40m track offset and 1.8±0.2 at 80m; injured PDEN was 3.4±0.4 at 80m.

Unlike roads and trails, neither distribution spiked within narrow track offsets, possibly due to subjects following easy near-stream travel paths. As with manmade linear features, PDEN variation by subject status was more apparent when discarding rescues and looking only at searches (Figure 14a).

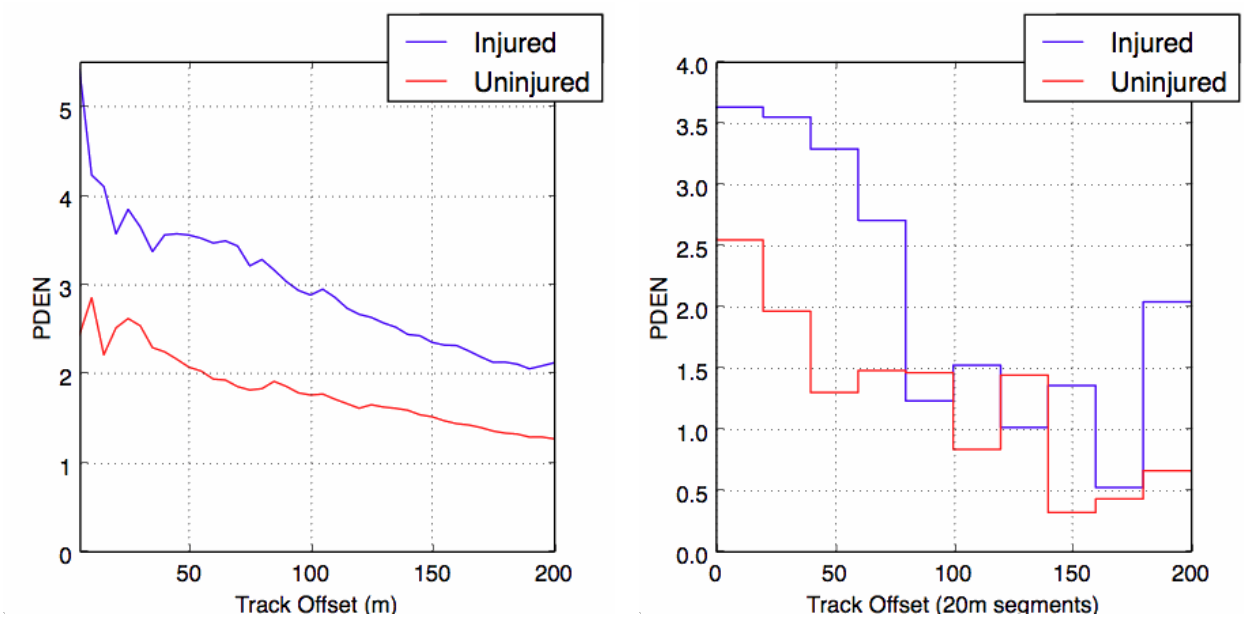


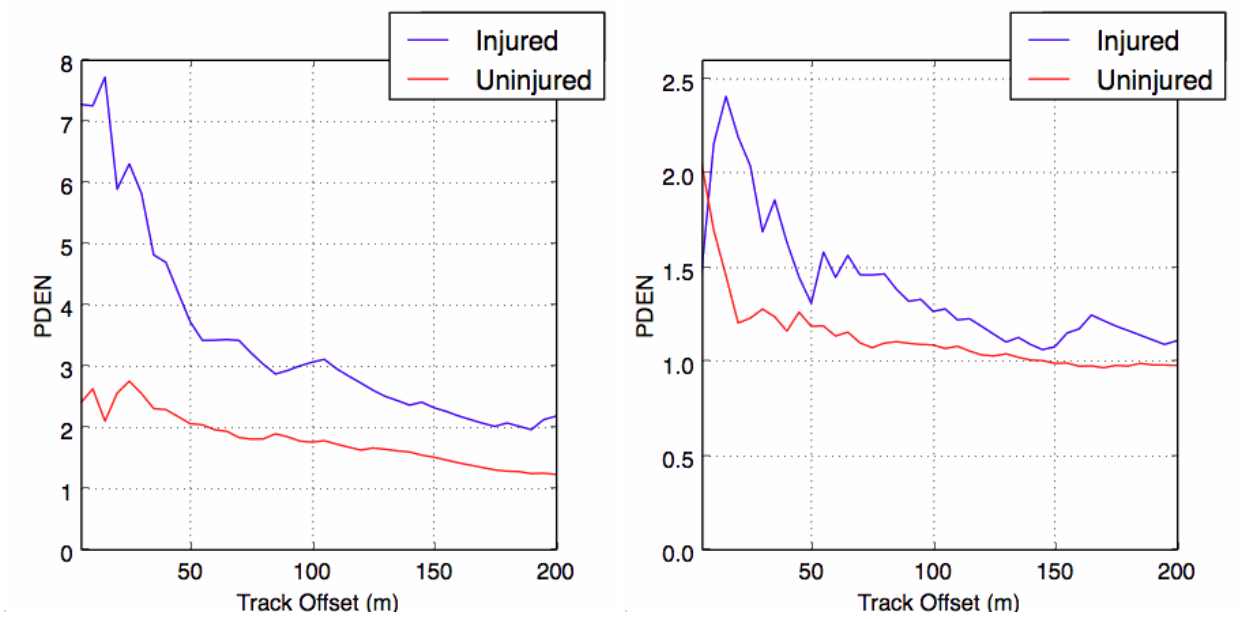
Figure 13: Stream PDEN.

Capillaries (flowlines with a Strahler number < 5) had a lower PDEN than streams, and given the number of near-capillary incidents that were also near-stream, were only marginally useful as a standalone metric. However, as with streams, their predictive ability improved when looking only at searches (Figure 14b). Further, when looking only at finds with a percentile basis elevation $>$ 0.2, i.e. mid-slope locations and above, capillaries and streams had similar PDEN.

4.3.3 Other Factors

Stream PDEN increased with vertical relief, although much of this increase is attributable to low-lying streams (Figure 15). Stream PDEN also decreased strongly with percentile basis elevation; low-lying streams were more predictive than mid-slope ones.

Stream PDEN may have increased mildly with IPP-find distance (Figure 16a), but the graph shown is hardly conclusive; only 40 near-stream locations had distances reported. While also based on a small number of incidents, capillary PDEN decreased with distance and capillaries had no predictive



(a) Streams *(b) Capillaries*

Figure 14: Cumulative water PDEN for search incidents only. The sharp rise in injured stream PDEN at track offsets less than 20m should be discounted due to the limited number of incidents sampled.

value beyond 3km. Taken together, these might suggest a trend towards more prominent features as subjects travelled farther from the IPP, but due to sample sizes, no solid conclusions can be drawn.

There was no statistically significant difference in stream PDEN between states. However, as streams accounted for fewer finds than roads and trails, only very large changes would have passed a significance test.

4.3.4 Summary

Interestingly, stream PDEN was similar for both on- off-trail finds. This made the intersection of trails (and to a lesser extent, roads) and streams particularly predictive; despite accounting for less than 1% of the search area, locations within 80 meters of both a trail and a stream comprised 7% of total finds for an uninjured PDEN of 7.5±1.4 and an injured PDEN of 13±3.

The predictive abilities of near-stream and near-manmade terrain were similar, although an exact comparison is difficult because positional error is greater for roads and trails than for streams. While subjects were more likely to be found directly on a manmade feature than a stream, high POD near-feature searching should give equal weight to streams and trails.

Subjects were far more likely to be found in low-lying streams than mid-slope ones. It's unclear whether this indicates an increase in PDEN with stream prominence, or if subjects simply headed downhill into canyon bottoms.

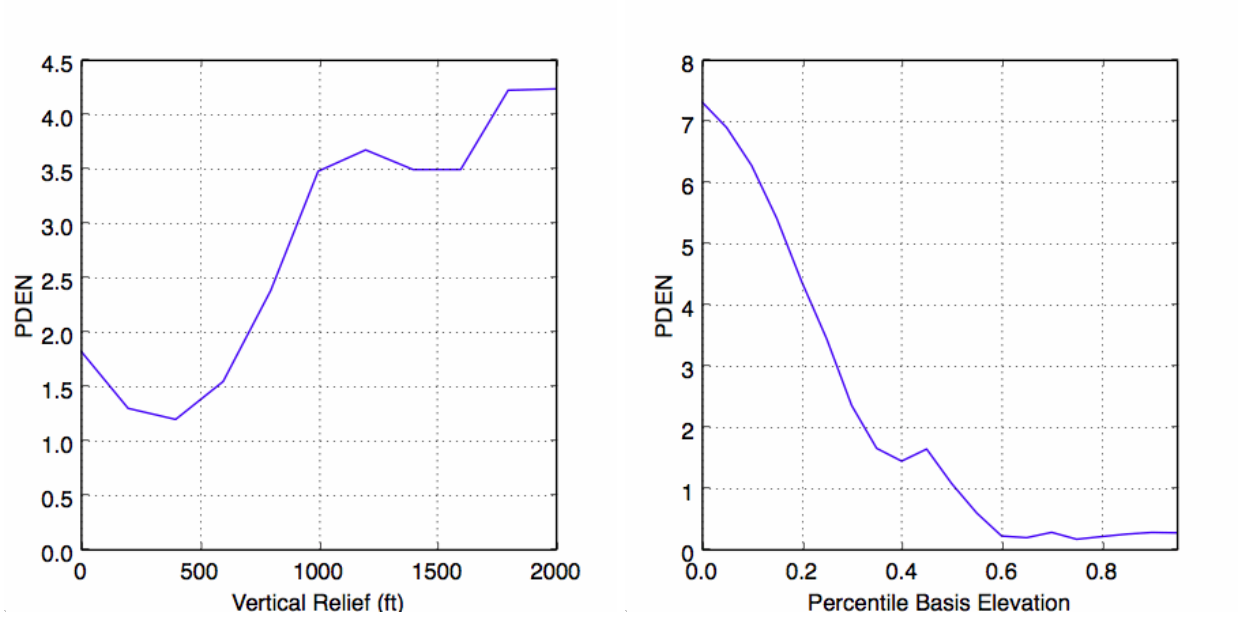
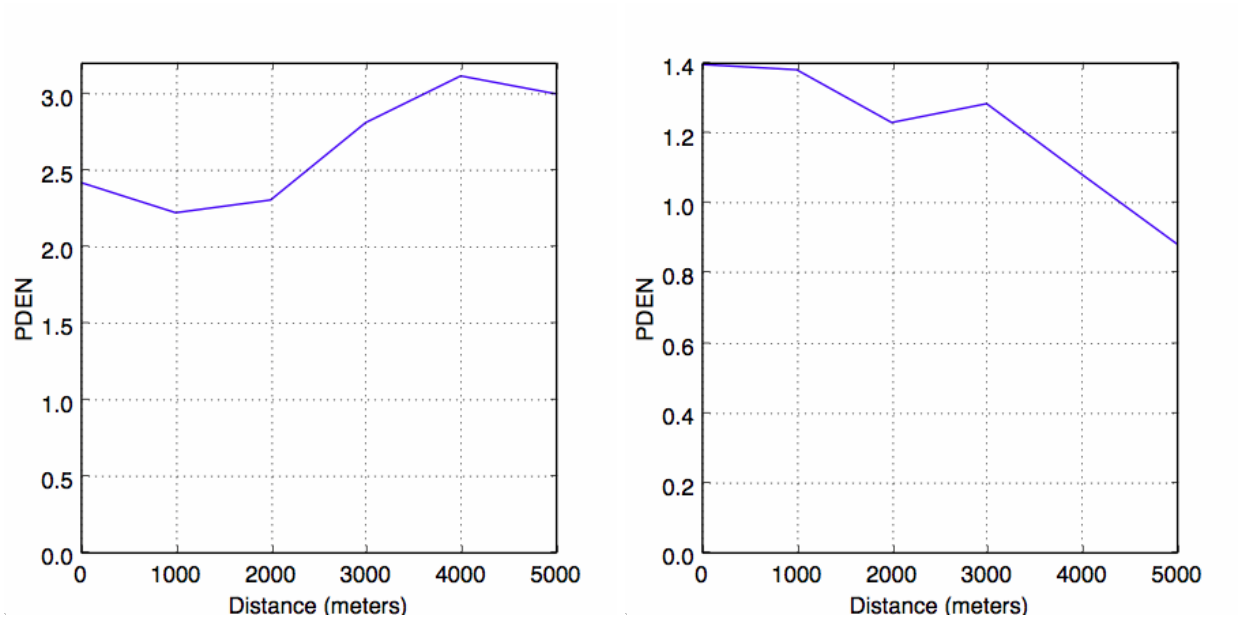


Figure 15: Effect of elevation on stream PDEN. Stream PDEN generally increased with vertical relief, although most of this was due to low-elevation streams, and decreased with percentile basis elevation.

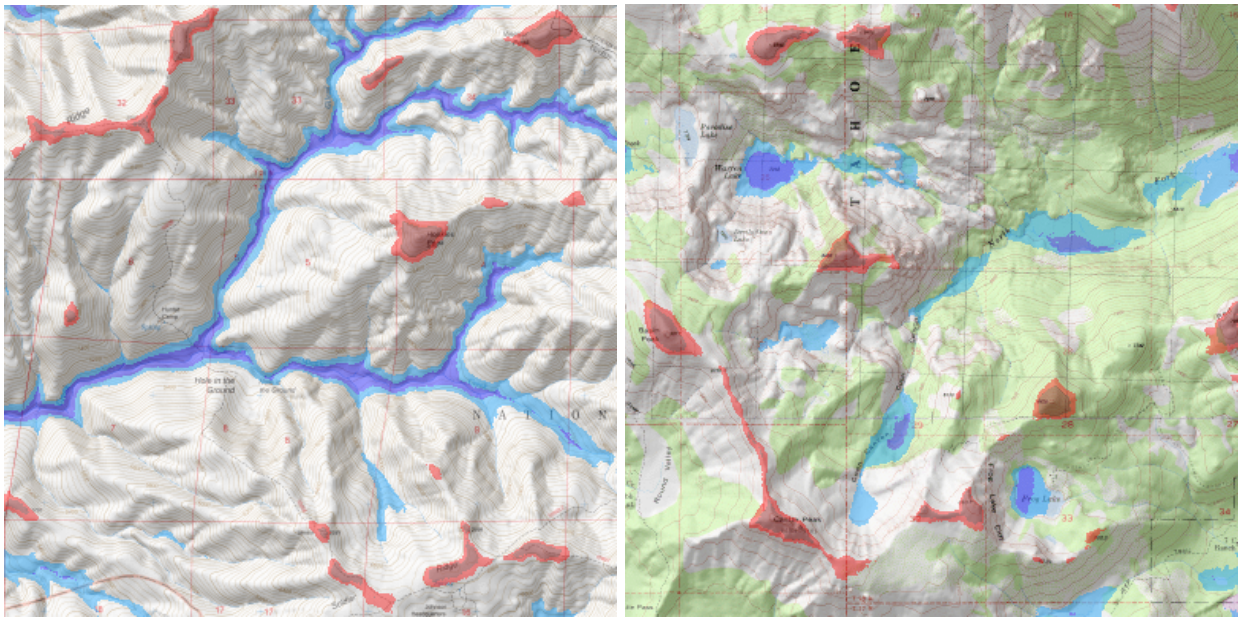


(a) Streams *(b) Capillaries*

Figure 16: Effect of IPP-find distance stream and capillary PDEN. Sample sizes were small at 40 streams and 50 capillaries.

4.4 Elevation

Find locations were evaluated using the National Elevation Dataset (NED). The NED provides elevation data for the contiguous US at 1/3 arc-second (approximately 10 meter) horizontal resolution, although in some areas this grid is interpolated from coarser data.



(a) Terrain with major river canyons. *(b) Terrain without major river canyons.*

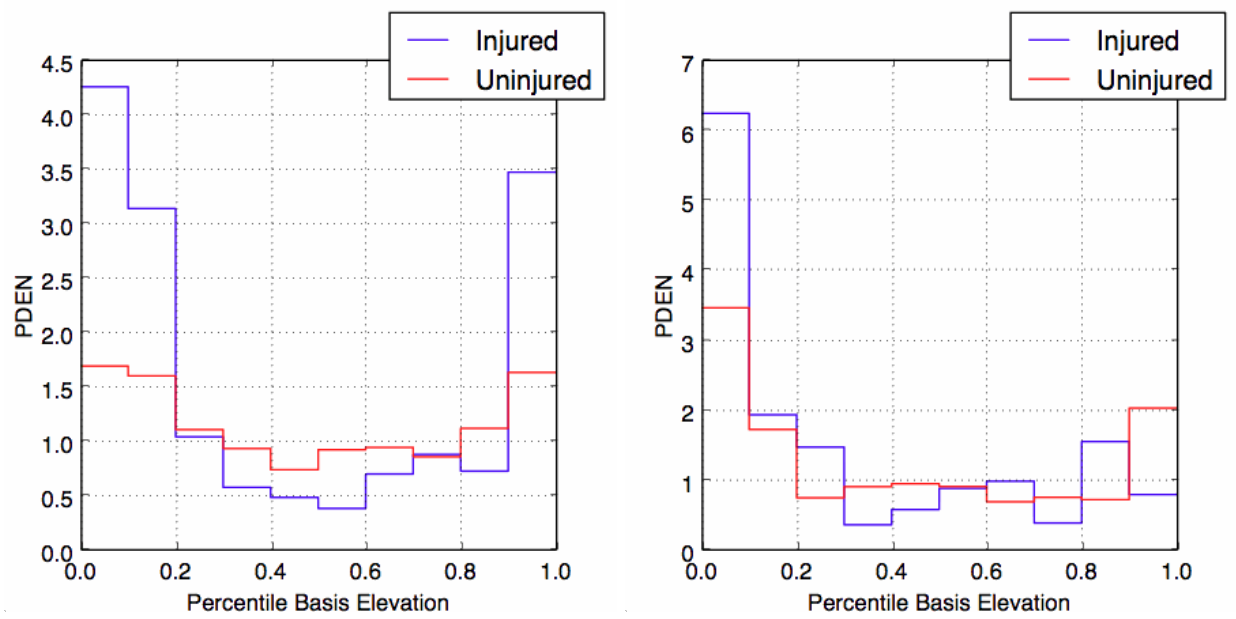
Figure 17: Example percentile basis elevations. Dark Blue < 0.1, Light Blue < 0.2. Light Red > 0.9, Dark Red > 0.95. PDEN increased in these areas, particularly at the darker-shaded extremes.

A segmented PDEN graph of percentile basis elevation (Figure 18) shows increased PDEN for both high and low points, with injured subjects having higher PDEN than uninjured ones. The error bars are quite large when subdividing by both subject status and the find location's proximity to a road or trail; the data is only subdivided this way in order to illustrate that elevation was predictive for both on and off trail finds.

The differences in low-elevation (< 0.2) PDEN between uninjured and injured subjects was statistically significant. Low elevation uninjured PDEN was 1.9±0.2 and injured PDEN 3.5±0.4; at extreme low points (< 0.1) injured PDEN was 5.3±1.

At high points, uninjured PDEN was 1.6±0.3 and injured PDEN 2.1±0.6. Both results were statistically significant, but the differences between them were not; it should be noted that the high elevation spike in the off-trail graph was due to only 10 incidents.

These PDEN numbers include all find locations, and rise when looking only at hilly and mountainous terrain. The predictive ability of low lying points for road and trail finds his suggests that lost, or at least injured, individuals will tend to head down into canyons and valleys even when on trails.



(a) Not on a manmade linear feature *(b) On a manmade linear feature*

Figure 18: Segmented percentile basis elevation PDEN, by find proximity to roads and trails.

4.4.1 Other Factors

PDEN for low-lying terrain generally increased with both vertical relief and IPP-find distance (Figure 19), although as with roads and trails, these relationships collapsed at the extremes. As one example, injured PDEN for extreme low points (< 0.1) rose from 2.6±1 in areas with $<$ 1000' of relief to 9.4±2 elsewhere. With at least 1000' of relief, those low points accounted for 2% of the search area but 20% of injured finds.

4.4.2 By State

Arizona did not have an increased high-elevation PDEN, although this may be a sample size issue and not a meaningful result. Regardless, it can not be applied to dry ecoregions as a whole; Oregon's dry and temperate ecoregions had similar high-elevation PDEN values.

4.4.3 Summary

Across both searches and rescues, and for both on and off trail finds, elevation was a strong predictor. Since low points have significant overlap with streams, it is impossible examine their effects independently, but low-lying canyon bottoms are worthy of extensive searching, even at great distances.

Figure 17 provides a real-world visualization of the high and low points that are worth targeting.

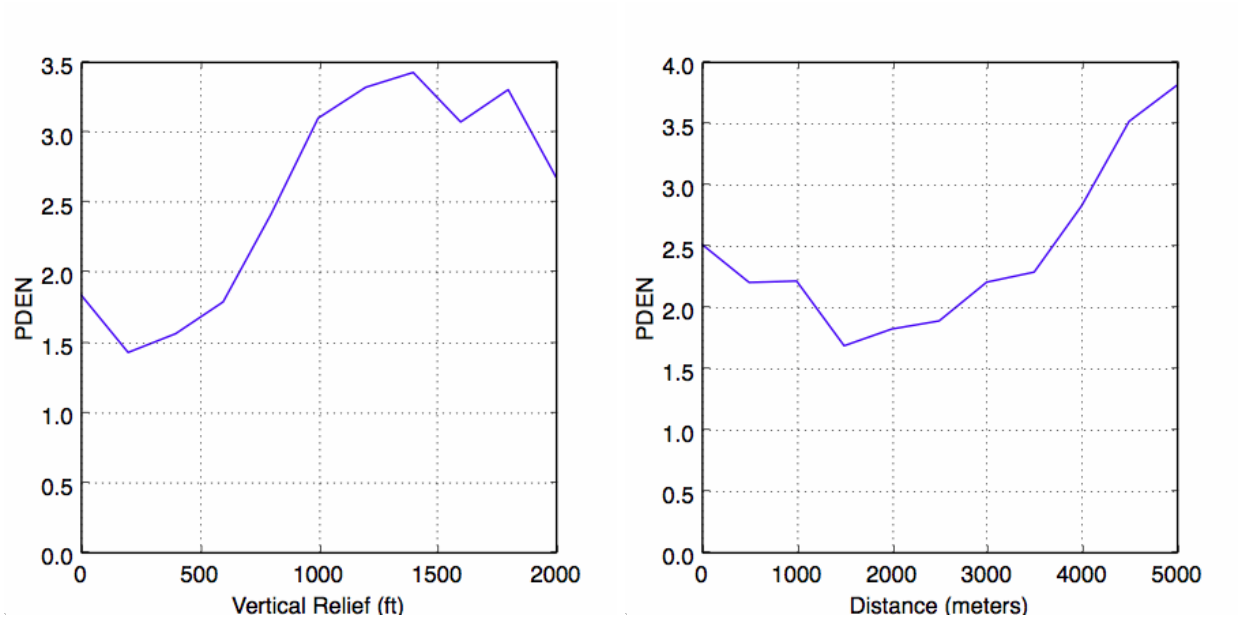
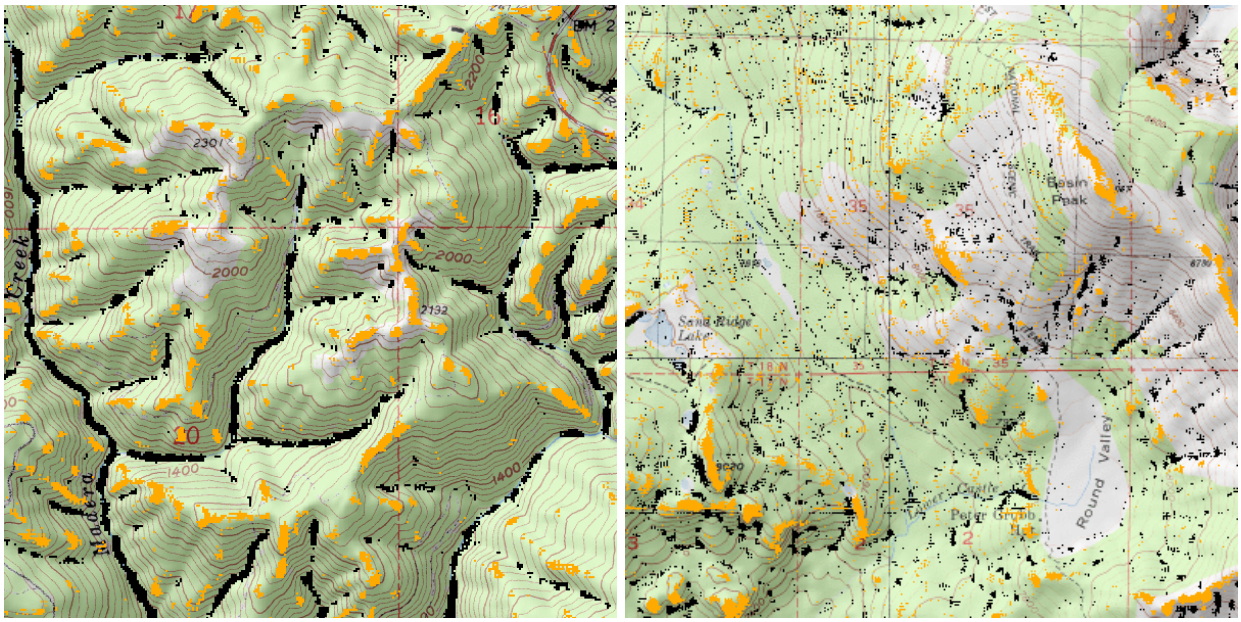


Figure 19: Effect of other factors on low elevation PDEN. PDEN generally increased with both relief and IPP-find distance.

4.5 Ridges and Drainages

While there is no dataset for ridges and drainages, they can be imperfectly modeled from elevation data using the Terrain Convergence Index (TCI) and Terrain Position Index (TPI). Both indicators only capture locations directly on convergent or divergent points, and not nearby finds. They may only capture intermittent locations along the linear ridges and drainages, and in areas without uninjured defined features, percentile basis TCI/TPI degrades to noise (Figure 20).

Because of these factors, ridge and drainage PDEN is likely understated, but it is difficult to guess by how much. More than any other feature, ridge and drainage PDEN is ripe for further investigation using new datasets or methods.



(a) Terrain with well defined drainages. *(b) Terrain without well defined drainages.*

Figure 20: Example percentile basis TCI. Orange < 0.05. Black > 0.95.

The Terrain Position Index (TPI) is the difference between a point's elevation and the average elevation of a surrounding grid. The result is a raw number, in units of length (e.g. feet or meters). The Terrain Convergence Index (TCI) also uses a grid, but measures the average dot product between each grid point's normal vector and the vector from grid point to sample point. The end result is somewhere between -1 (complete divergence) and 1 (complete convergence). The TCI in this paper uses a 3-D vector incorporating elevation and slope angle.

Because the tiled elevation data used for this project is stored as integer meters and experiences discrete jumps, the 3x3 grid usually used for TCI and TPI resulted in excess noise in flatter terrain. To address this, TCI and TPI were computed using a 5x5 grid with additional smoothing applied.

4.5.1 Off Trail Finds

TCI and TPI were only examined for off-trail finds.

Percentile basis TCI and TPI both had predictive value, but were weaker than elevation (Figure 21). Although the graphs are presented with a segment size of 0.1 for visual clarity, most of the PDEN increase was due to to the top and bottom 5% (percentile basis < 0.05 and > 0.95), each about 5% of the total search area.

For injured subjects, convergent terrain (percentile basis TCI > 0.95) had a PDEN of 2.9±0.6. The same effect was observed when looking at raw TCI instead of percentile basis; TCI > 0.05 was 5% of the search area but 15% of injured finds.

For uninjured subjects, convergent terrain had a PDEN of 1.4±0.3. Divergent terrain (percentile basis TCI < 0.05) had a PDEN of 1.8±0.4. While both of these effects are mild, they were still found to be statistically significant ($p = .04$ and $p < .01$ respectively)

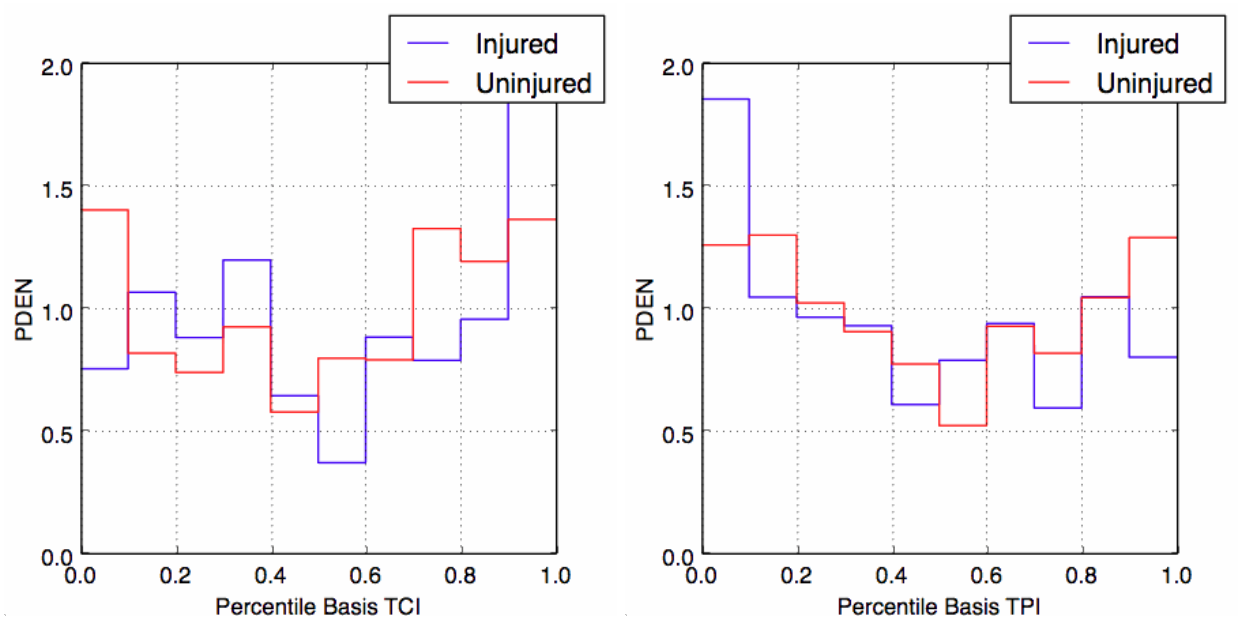


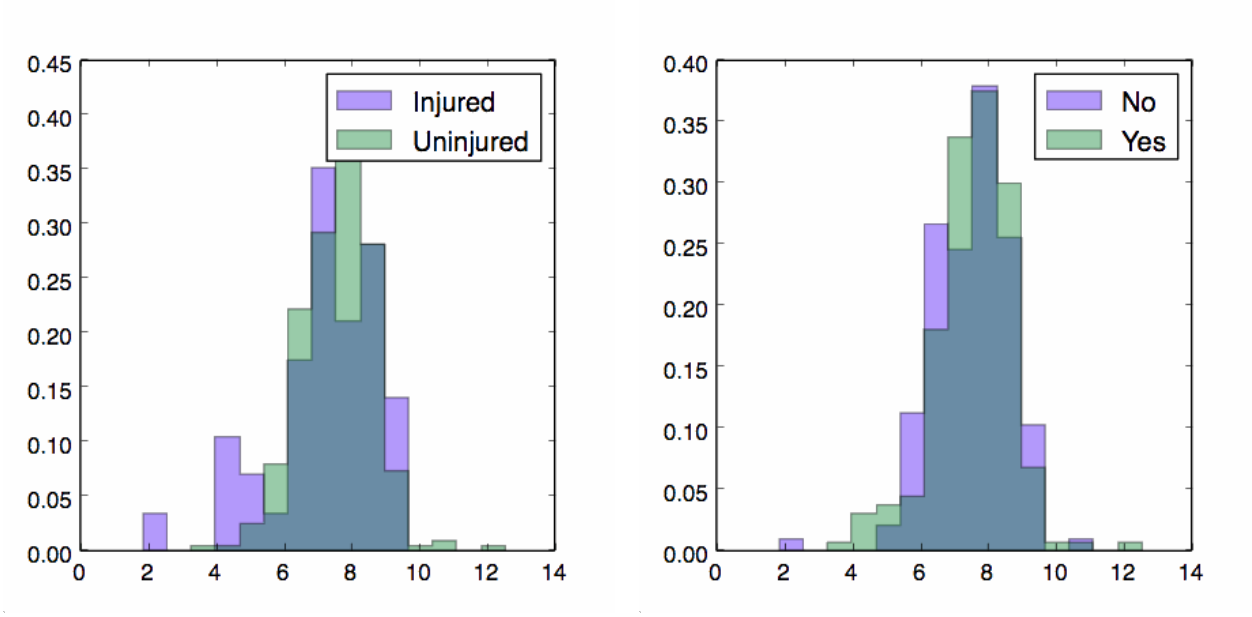
Figure 21: Percentile basis convergence PDEN for off-trail incidents. The x-axes mirror each other; convergence is associated with high TCI (left) and low TPI (right).

4.5.2 Other Factors

Although summarized here as "ridges and drainages", convergent terrain also includes low points like canyon bottoms, and divergent terrain includes high points and benches. Due to the sample sizes involved, it was not possible to fully separate the effects of convergence and elevation. It was also not possible to meaningfully examine how TCI PDEN varied by IPP-find distance and vertical relief.

4.6 Additional Influences

4.6.1 Distance



(a) By subject status

(b) By IPP proximity to manmade linear features

Figure 22: Log IPP-find distance histograms. Subject status and the IPP's proximity to a manmade linear feature both had only mild effect on IPP-find distance.

Surprisingly, IPP-find distance was only weakly influenced by subject status (Figure 22a). Ignoring reported distances below 10 meters, the median IPP-find distance was 2.2km for uninjured subjects and 1.7km for injured ones. Similarly, IPP-find distance was only weakly influenced by the IPP's proximity to a mapped road or trail (Figure 22b).

4.6.2 Land Cover

The National Land Cover Database (NLCD) provides surface coverage information for the contiguous US at a 30 meter horizontal resolution. Each grid cell is assigned a coverage type such as forest, grassland, barren, water or developed. The NLCD USFS Tree Canopy Cartographic product supplements this with tree canopy coverage at the same 30 meter horizontal resolution; a canopy value of 100 means that 100% of the grid cell is covered by tree canopy.

Despite expectations that clearings or thickets would have predictive value, canopy coverage and land cover were not found to affect PDEN. This may be due to the techniques used; one unexplored avenue is to look at interfaces such as the edges of clearings or marshes.

PDEN did increase in areas with low canopy coverage, but this was due to roads registering on the canopy and NLCD datasets and was not discernible in off-road finds.

4.6.3 Slope Angle

Off-trail rescues had an increased PDEN for steep slopes, but searches did not. For off-trail rescues, percentile basis slope angles > 0.9 had a PDEN of 1.7 ± 0.35 and slope angles $> 35°$ had a PDEN of of 2.3 ± 0.4;

4.6.4 Track Offset

At first glance, the relationship between PDEN and track offset presented here seems difficult to reconcile with the track offset tables in *Lost Person Behavior*. While these findings show little benefit in searching beyond 100 meters of a linear feature, Koester advises in *Lost Person Behavior* that "the fact that 95% of despondents are found within 500m of a linear feature is important considering that the maximum zone for despondents is over 13.3 miles from the IPP" (p 86). Does it make sense to create broad area segments centered on linear features, or not?

Consider hikers - *Lost Person Behavior* lists 50% off hikers being found within a 100m track offset and 95% within 424m; the types of linear features measured against are not specified. A logical interpretation of these numbers is that after initial sweeps fail to turn up the subject, a search should be focused on areas within 400m of linear features.

A different perspective is that within the search areas studied in this paper, a surprisingly large amount of terrain was found to be near linear features. Looking only at roads, trails and streams, 35% of the search area was within a 100m track offset, 42% between 100m and 400m, and 23% beyond 400m. The 100m-400m zone had roughly twice as many finds as the $>$ 400m zone (22% v. 12%), but since it also had twice the area, their PDENs were equal.

Barring more information about the linear features that ISRID track offsets are measured against, there is no meaningful way to compare ISRID's 45% of finds against this paper's 42% of terrain. However it seems likely that the track offsets presented in *Lost Person Behavior* are not, by themselves, sufficient to determine find probabilities. While useful for comparisons between subject categories, the raw percentile offsets should not be used to determine near-feature segment sizes.

4.6.5 Other Factors

There were no meaningful differences found between hikers, hunters and gatherers. A greater percentage of hikers were found near trails than hunters, mirroring conventional wisdom that hikers are more trail oriented. However, they were also found in areas with greater trail density; accounting for search area composition, PDEN was equal. Although difficult to separate from other factors, no patterns were found relating terrain based PDEN to age or sex.

Elevation change from the IPP was not found to be a useful predictor beyond its established use in distance based models. However, since elevation change increased with distance, median elevation gain/loss was 6 times greater at distances beyond 2km than distances within 2km. This calls into question the validity of predicting subject behavior using raw historical elevation changes. A possible alternative could be to use glide ratio, i.e. vertical change divided by horizontal change. Median downhill glide ratio was less dependent on IPP-find distance, at 6% within 2km and 4% beyond.

5 Conclusion

5.1 Behavior Summary

There is no perfect way to distill these results into a single set of numbers. PDEN increases near linear features and at the extremes of terrain, and restricting searches to a handful of high-probability areas leads to impressive multipliers. Broadening the scope to include more finds causes the numbers, while still elevated, to drop. Positional error is greater for trails than roads, and roads than streams, making direct comparisons between features difficult.

Table 1 is my best attempt to provide generic PDEN numbers, but it should be used as a rough guide only, and not as the basis for detailed calculations. Remember that injured subjects are comparatively more likely to be near features, rather than on them. Many linear features are worth searching to track offsets greater than a single sweep can reasonably achieve. An initial narrow-offset search might have a higher PDEN than shown, while a wider-offset followup later in the operation would have a lower one.

Feature PDEN generally increases with both IPP-find distance and the vertical relief of surrounding terrain. As just one example, roads and trails increase from 30% of finds within 2km of the IPP to 53% beyond 3km However, until additional data points allow for more fine-grained stratification, there is no way to provide hard PDEN numbers for specific distances or terrain types.

Feature	**Uninjured**	**Injured**
Roads	3x	1.5x
Trails	5x	7x
Lakes	2x	2x
Streams	2x	3.5x
Capillary Streams	–	1.5x
Stream / Trail Interfaces	7x	12x
Low Points	2x	5x
High Points	1.5x	2x
Ridges	2x	–
Drainges	1.5x	3x

Table 1: Probability effect of various terrain features. Road and trail PDENs in particular will be higher during hasty searches and lower when searching nearby.

5.2 Major Findings

1. Injured and uninjured subjects have different probability distributions, allowing searches to employ a targeted mix of high and low POD searching.

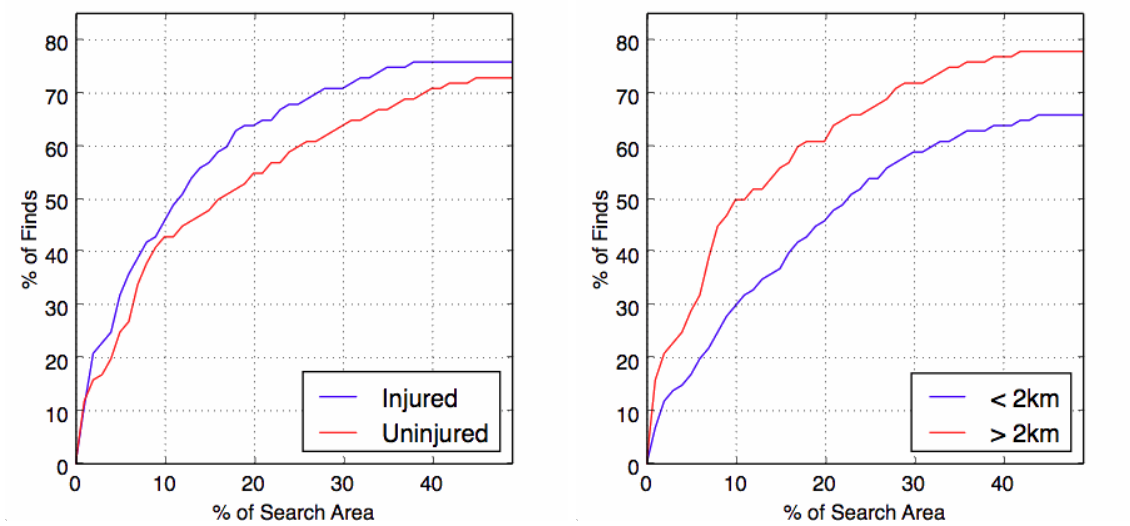
2. Gridded area searching should only be conducted after all linear features within the area have been exhausted.

3. For high POD (i.e. not hasty) search efforts, streams and drainages should be given equal weight to roads and trails.

4. Roads, trails and streams should be searched to a maximum track offset of approximately 100m, although smaller offsets may be warranted. Focusing search efforts on areas within 1/4 or 1/2 mile of a road or trail network does not appear to be an effective strategy.

5. Search efforts can be increasingly focused on major linear features as distance from the IPP increases. Careful consideration should be given before limiting long distance searches of those features based on behavioral statistics.

6. There is no magic bullet. Terrain based models provide some quick-win strategies, but ultimately only 60%-80% of subjects are found on high points, low points, ridges, drainages or within 100m offsets from roads, trails or streams. While that initially sounds impressive, those features account for 45% of the search area (Figure 23).



(a) By subject status *(b) By IPP-find distance*

Figure 23: Percent of finds compared to percent of search area covered, using terrain based models. These graphs are for illustration purposes only and do not represent a progressively executable search plan. At each spot on the x axis, a new mix of features was chosen that maximized the number of finds for the given search area coverage.

5.3 An Integrated Approach to Search Management

The findings presented here are most useful if they can be distilled into an actionable search plan. Prescribing a new search management model is well beyond the scope of this paper, but I'd like to show how terrain based probability models could be applied to current search management techniques.

The classic bicycle wheel model defines a hard limit on the search area (the rim), and a two-step process of hasty searching linear features (the spokes) followed by gradually filling in the remainder with area searching. The rim can be expanded as interior search coverage increases, but at this point the search is generally presumed to be for an unresponsive subject.

My interpretation of the evidence is that this model should be adapted into a spiderweb: dense searching at the core surrounded by a web of linear features, with several long-ranging stringers at the periphery. Table 2 provides a high level structure for the web, but its exact form will inevitably vary across searches. Among other factors, an IPP's predictive ability will influence the proper balance of distance and terrain based probability models.

Zone	**Tactic**	**Target**
Stringers	Responsive	Manmade Linear Features, Low Terrain
	Unresponsive	Stream / Trail Interfaces
Outer Web	Responsive	Natural Linear Features, High Terrain
	Unresponsive	Low Terrain
Inner Web	Responsive	Areas
	Unresponsive	Linear Features, High Terrain
Core	Unresponsive	Areas

Table 2: Spiderweb model search tactics, ordered by distance from IPP.

Rather than comprehensively checking off features in order, evidence supports expanding ongoing searches by growing the entire web. As gridded area searching is extended farther out from the IPP, high-POD searching of roads, trails and drainages should be similarly extended to greater distances. As compared to the classic model, a broader mix of targeted search tactics remain deployed throughout an operation.

The spiderweb model presents added challenges to probability consensus development. From a big picture perspective, terrain based micro-segmentation results in an unmanageable number of segments for a Mattson's style process. This necessitates a distinction between larger probability regions and smaller search segments, which while not a new concept, is not always practiced.

At small scales, terrain based models require building consensus around questions that traditional processes are ill-equipped to answer. A simple example would be the likelihood of a subject intentionally leaving a trail, which is an important consideration in determining trail PDEN. Successful application of a spiderweb model may require new processes for forming consensus around terrain-based probability changes.

6 References

6.1 Further Reading

Hill, K., & O'Connnor, D. (Eds.). (2007). *Managing the Lost Person Incident*
National Association for Search and Rescue (NASAR)

Stoffel, R. (2006). *Textbook for Managing Land Search Operations.*
ERI Publications & Training

Koester, R. (2008). *Lost Person Behavior*
dbS Productions LLC

Cooper, Frost, Robe (2003). *Compatibility of Land SAR Procedures with Search Theory*
http://www.uscg.mil/hq/cg5/cg534/nsarc/LandSearchMethodsReview.pdf

Koester, R., Cooper, D.C., Frost, J.R., Robe, R.Q. (2004). *Sweep Width Estimation for Ground Search and Rescue*
http://www.uscg.mil/hq/cg5/cg534/nsarc/DetExpReport_2004_final_s.pdf

Doke, J. (2012). *Analysis of Search Incidents and Lost Person Behavior in Yosemite National Park*
http://kuscholarworks.ku.edu/bitstream/handle/1808/10846/Doke_ku_0099M_12509_DATA_1.pdf

6.2 Data Sources

International Search and Rescue Incident Database
http://www.dbs-sar.com/SAR_Research/ISRID.htm

OpenStreetMap
http://www.openstreetmap.org/

FSTopo Transportation Dataset
http://data.fs.usda.gov/geodata/vector/index.php

National Elevation Dataset
http://ned.usgs.gov/

National Hydrography Dataset
http://nhd.usgs.gov/

2011 National Land Cover Dataset
http://www.mrlc.gov/nlcd2011.php