JSAR

JOURNAL OF SEARCH & RESCUE

ISSN 2230-5734

JournalofSAR.com

VOLUME 6

ISSUE 1

Volume 6, Issue 1, January 2023
Pages 1-75
ISSN 2230-5734

Journal of Search & Rescue

The Journal of Search and Rescue (JSAR) is an open access peer-reviewed electronic journal for the collation and distribution of original scholarly material on search and rescue (SAR).

It is being supported by the in-kind work and contributions of the Editorial Board. There is still the need for a dedicated journal serving those with a direct interest in all disciplines of search and rescue including: rope rescue, water (flat, swift and marine), ice rescue, wilderness search and rescue, structural collapse rescue, trench collapse rescue, cave rescue, dive rescue, motor vehicle extrication, canine search, technical animal rescue, air rescue, search theory, search management, and mines rescue. JSAR exists to fulfil that need.

Article submissions from these and other SAR disciplines are welcome. Launching this journal on the internet offers a relatively cost-effective means of sharing this invaluable content. It affords the prompt publication of articles and the dissemination of information to those with an interest in SAR.

JSAR will provide a forum for the publication of original research, reviews and commentaries which will consolidate and expand the theoretical and professional basis of the area. The Journal is interested in theoretical, strategic, tactical, operational and technical matters.

Advertising within JSAR will be considered in the future to ensure sustainable funding is available to enhance and continue the work of the journal. The publication of an article in the Journal of Search and Rescue does not necessarily imply that JSAR or its Editorial Board accepts or endorses the views or opinions expressed in it.

Cover Image: Photo by Scott C. Hammond. Utah County Sheriff's Search and Rescue, USA

Volume 6, Issue 1, January 2023
Pages 1-75
ISSN 2230-5734

Journal of
Search & Rescue

Contents

Volume 6, Issue 1, January 2023
Pages 1-75
ISSN 2230-5734

Journal of
Search & Rescue

In memory of LCDR Frederick Koester 1932-2022

Volume 6, Issue 1, January 2023
Pages 1-75
ISSN 2230-5734

Journal of
Search & Rescue

Editorial

Welcome to Issue 1 of the 6th Volume of the Journal of Search and Rescue. The recognition of the importance of SAR is increasing, as natural disasters increase in their frequency and ferocity. The Intergovernmental Panel on Climate Change (IPCC) have stated that extreme weather events and their impacts will increase throughout the 21st Century, and clearly this will impact upon the SAR community in both their work and lives.

It is essential that SAR organisations utilise methods and technology to optimise the effectiveness and efficiency of their operations through research and sharing best practices. The authors in this edition of JSAR utilise case studies, existing data sets and experimental techniques to identify strengths and weaknesses in various SAR techniques.

The dreaded ‘networking’ can often result in the development of ideas and SAR conferences are a perfect place for this. Lead JSAR editor, Dr Robert Koester, presented at the Icelandic SAR conference, Rescue 2022, and returned laden with business cards and prospective meetings. The Institute of Search and Technical Rescue hosted their inaugural annual conference in November 2022. Taking place at the Union Jack Club, London, the conference brought together experts in the field of SAR to share examples of past, present and future emergency response events, in order to develop SAR practices. Talks included topics on academic assessments of SAR using drones; how to improve cohesion, recruitment and selection in SAR teams; sea rescue of migrant crossings in the Mediterranean; the importance of putting the casualty at the centre of rescue efforts; and the risks/benefits of immersive training. The keynote speaker was Rick Stanton who gave a fascinating talk on his role in the Tham Luang cave rescue. The lessons learnt and shared were inspiring and enabled practitioners from across the world to forge new relationships to further their SAR skills.

International collaboration provides unique opportunities to train and be trained, building capacity for both home and visiting SAR teams. ServeOn, the UK lead for EVOLSAR (European Association of Civil Protection Volunteer Teams), have been invited to participate in a large-scale training event in Kosovo in March 2023. In the past six months UKISAR have been deployed for training in Kosovo, Singapore and Switzerland. As well as capacity building this is preparing UKISAR for the UN INSARAG External Reclassification in 2023. And on that note congratulations to SARAID (Search and Rescue Assistance in Disasters) for becoming the first and only UK voluntary team to be officially classified by the UN INSARAG as a Light USAR Team and only the third worldwide.

Lessons can be learnt from successes and failures and one such lesson is on the subject of equality, diversity and inclusivity. Following the November reports on alleged incidents of misogyny, racism and bullying in the London Fire Brigade* we should all reflect how SAR organisations, either professional or voluntary, can foster an inclusive environment to recruit, and retain, a diverse range of people. There are many fantastic examples of inclusivity in the sector but we should all be striving to do better to access the skills of underrepresented people. We must question our unconscious biases and representation in recruitment. What actions, languages and behaviour could impact on retainment. There have been reports of women joining voluntary SAR teams only to find that none of the PPE equipment is suitable for their body size and shape. SAR teams should reflect and respond to the communities they serve, and diversity brings strengths and attributes that enable us to achieve that aim more effectively.

*(*JSAR has no position on, nor makes any comment on the allegations themselves)*

Toby Meredith, Portsmouth UK

Deriving Objective Probability of Detection for Missing-Person Search: Validating Use of Effective Sweep Width and Associated Mathematical Models

Kenneth B. Chiacchia, PhD

Heather E. Houlahan, AB

Mountaineer Area Rescue Group

Appalachian Search and Rescue Conference

Morgantown, WV

U.S.A.

Email kchiacchia@asrc.net

Abstract

Effective sweep width (*W*) promises objective probability of detection (POD) values for guiding missing-person search efforts. However, methods for measuring *W* produce large uncertainties. Also, models for generating POD from *W* have not been validated for ground-based search. The authors applied least-square fits of POD data collected in the field for air-scent dog teams as well as human searchers using the two most prevalent models to derive *W* values. The method routinely fits the data to an R-square of >.8 with more statistical power than previous methods, and appears to be detector-agnostic. The authors present recommendations for optimizing its use in the field.

KEYWORDS: SAR, effective sweep width, probability of detection, search theory, dogs, smell

Introduction

In ground-based search for missing persons, resource allocation is a major imperative. Available resources — particularly trained searchers and specialized search resources — are virtually never sufficient to look everywhere the search subject could potentially be. Thus, operational search managers employ Bayesian logic to prioritize which "segments" — fractions of a given search area small enough to be searched by a single team/modality — will be searched at a given time via the equation:

1. $POS_o = \sum(POD \times POA)$

where POS_o is overall probability of success for the entire search effort (*i.e.*, finding the subject; Koopman BO, 1980; Charnes A, Cooper WW, 1958). POD, probability of detection, is the probability for a given detector/group of detectors to identify a given search object in a segment under given environmental conditions, assuming the object is located in that segment. POA, or probability of area, is the probability that the desired search object *is* contained within that segment.

By summing POD X POA (*i.e.*, the POS for each segment) over all segments searched to date, a current estimate for POS_o can be derived, because each unsuccessful search effort (*i.e.*, not generating a find) in a segment reduces its POA. (The new POA is [1 – POD] X the initial POA.) This approach — sometimes called "shifting POAs" — continually reassesses POAs throughout the search area, allowing search managers to reallocate efforts toward the areas with the highest POAs (weighted by time required to search), thus optimizing POS_o (Charnes A, Cooper WW, 1958). Shifting POAs can thus be used to determine whether to continue searching current segments, expand to a larger search area, or indeed to suspend the search pending better information on the subject's possible location.

Much work has been done to amass historical data to estimate initial POAs (Syrotuck WG, Syrotuck JA, 2000; Koester RJ, 2008; Koester RJ, 2020) with some confidence. On the other hand, objectively measuring POD, given the likely effects of different detectors and environmental factors, remains an unfinished project within the search-and-rescue (SAR) community. An expanding body of work has focused on obtaining objective PODs through field measurement of effective sweep width, or *W* (the distance defining an envelope within which a given combination of detector/search object/conditions experiences an equal number of detections and misses; Cooper DC *et al.*, 2003; Koester RJ *et al.*, 2004; Chiacchia KB, Houlahan HE, 2010; Koester RJ *et al.*, 2014; Chiacchia KB, 2020). New mapping/search planning computer applications are beginning to incorporate *W* values. The old practice of determining POD through subjective estimation by search-team leaders, which produces highly inaccurate values (Koester RJ *et al.*, 2004), may still be more common than use of *W*; but the SAR community is making progress.

Issues remain in the use of *W* to derive objective POD values. One study involving computer modeling and field measurements suggested that navigational errors in search patterns may impose errors of ~10-20% in prospective POD determinations of intended search tasks (Perkins D, 2018) — though use of GPS tracks for searchers can document the actual search pattern and so avoid this issue in retrospective assessment of already-completed efforts. In addition, use of *W* almost certainly requires adoption of computerized search planning, since the nonlinear relationship between *W* and POD is unintuitive (Koopman BO, 1980; also, see Results and Discussion).

A more significant problem may derive from the limits of field-measurement of *W*. *W* values can be derived from the canonical crossover method (Koopman BO, 1980; Cooper DC *et al.*, 2003; Koester RJ *et al.*, 2004), which determines the "crossover point" at which the number of detections, cumulated from infinite distance inward, equals the number of misses, cumulated from zero distance outward. The lateral distance of the crossover point is *W*/2, as the detector "sweeps" this distance to its left and its right as it moves through the search area. By making a crossover plot for each individual searcher or dog-and-handler team, we can derive mean *W* values and statistical parameters for those means. However, the small data from each individual determination tends to result in high uncertainties, limiting us to only modest statistical comparisons of detectors and effectors of *W* (Chiacchia KB, Houlahan HE, 2010; Chiacchia KB *et al.*, 2015).

An examination of a typical crossover plot, seen in Figure 1 (adapted from Chiacchia KB *et al.*, 2015), explains part of the problem. Again, the plot cumulates detections from infinity and misses from zero distance, with the crossover point defining *W*/2. In this particular case, generating the crossover plot involved collection of more than 250 detection opportunities (DOs, lateral detections or misses). However, a simple thought experiment shows that adding detections below the crossover point or misses above it does not change the crossover point and thus *W*. The value for *W* rests solely on detections at the crossover and above and misses at the crossover and below — again in this case, only slightly more than 50 out of 250 DOs "count" toward *W*, a significant loss of data.

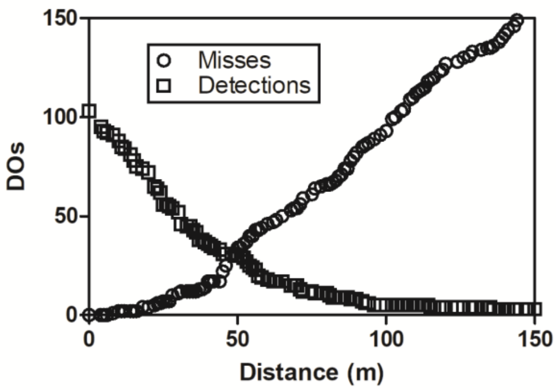


Figure 1: Crossover plot for air-scent dog teams, adapted from the authors' earlier report. Note that, as detections are cumulated from infinity and misses cumulated from zero lateral distance, only detections ≥ the crossover distance or misses ≤ the crossover distance actually contribute to determining W (which is twice the crossover distance). Reprinted from Wilderness & Environmental Medicine*, 26(2), Chiacchia KB* et al.*, Deriving Effective Sweep Width for Air-scent Dog Teams, p. 146, Copyright (2015), with permission from Elsevier.*

An alternative method, the lateral range curve, can be used to determine *W* and makes more complete use of the data (Koopman BO, 1980; Cooper DC *et al.*, 2003; Koester RJ *et al.*, 2004). But this method depends on calculating detections at each given lateral distance from a detector, which either means the huge noise of low-DO (<~30) determinations at each distance (sometimes an impossibility, when there is only a single DO at that distance) or binning of distances, which adds averaging bias. See Figure 2 for an example of this problem in a typical lateral range curve, generated by the data for the low-visibility mannequin in summer in our earlier report (Chiacchia KB, Houlahan HE, 2010). (As POD is by definition the probability of detection within a given area, the detections here, being single-distance probabilities, are instead labeled "% Detected;" Koopman BO, 1980)

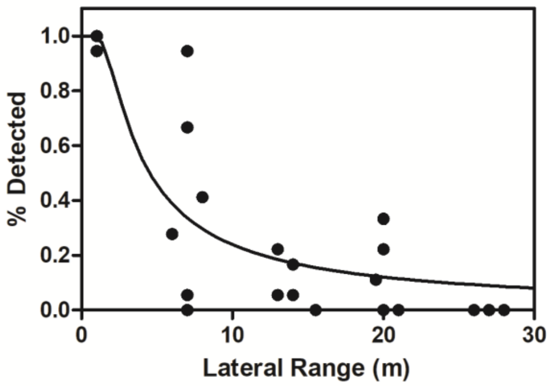


Figure 2: Typical lateral range curve, from the low-visibility mannequin in State Game Lands 203, summer (Chiacchia KB, Houlahan HE, 2010), showing the high degree of noise in individual points based on low number of DOs. The y-axis is labeled "% Detected," as single-distance detections do not fit the area-based definition of POD. The fit line, for illustration of the trend, is arbitrarily based on the inverse-cube model (see below).

Another problem with determining objective PODs from *W* is that the predominantly used "random-search" model (a name referring to random error in the detector's intended path, not a detector searching a segment in a random path) for deriving POD from *W*, while used successfully for some time in the maritime- and aerial-search spheres (Koopman BO, 1980), has never been explicitly validated in ground-based search. Hence, some SAR researchers and practitioners are skeptical that the method produces accurate results in that sphere.

In this report, the authors present a new way of determining *W* that explicitly links the value to field-measured PODs for both air-scent dog teams and human visual searchers. Unlike the crossover method, it employs virtually all the data collected; and without binning it produces much less noise in individual points in the curve than the lateral range method. The more efficient use of data helps to enable better least-square fits and thus more extensive statistical comparison of detectors and effectors of *W*, as well as testing of the appropriateness of specific theoretical models for the relationship between POD and *W*.

Methods

Air-Scent Dog Team Experiments

This report analyzes data collected anew, between 2016 and 2020, as well as data from our previous report (Chiacchia KB *et al.*, 2015). We collected new air-scent dog team data using the methodology described in that report. Briefly, we recorded GPS data on randomized, blinded daylight air-scent tasks of 5 ha or larger (to ensure the possibility of longer-range finds and misses), using a Garmin GPSMap 60csx or 62sc GPS receiver (Garmin International, Olathe, KS, USA), tracking the handler's path and recording the position at which the dog signaled a find to the handler and the position of the subject when the dog led the handler in to the subject. No finds by the dog without such a "refind" that successfully led the handler in to the subject were scored as "detections;" lateral approaches to the subject without such a successful refind, whether misses or detections without a refind, were scored as misses. As reported, all detection or miss distances were recorded laterally from the handler's track and not directly between the handler's position at the refind and the subject, in some cases requiring extrapolation of the handler's intended path. Detections and misses were recorded manually using a 1:3,000 or 1:4,000-scale printout of the task using either MyTopo Terrain Navigator Pro (various versions, MyTopo, Billings, MT, USA) or the online version of SARTopo (CalTopo LLC, Truckee CA, USA).

For the current report we collected DOs from 178 air-scent dog tasks at 27 locations throughout western PA and northern WV, USA. Of these tasks, 11 were excluded from the study: 1 lacked a defined search area in our records, 2 because noticeably poor performance by the dog was followed by a diagnosis requiring Lyme disease treatment, 2 because the subject was not inside the assigned search area, 3 because the subject inadvertently unblinded the problem, 1 because fencing not reflected on the USGS maps prevented the team from fully accessing the assigned area, 1 because the team searched the wrong area, and 1 because the subject was in motion during the task and so relative position with the handler's track was not possible to obtain. The remaining tasks generated 150 detection and 393 miss distances, a total of 543 DOs. Note that when the team completed the task without a find (which happened in 10 tasks), the problem was unblinded so that the dog could make a find for training purposes and only misses scored before the unblinding were counted. As with the earlier study, and because it is difficult to distinguish olfactory finds by the dogs and (relatively rare) visual finds by the handler, both were scored as "dog team finds," with the difference in *W*s between dog teams and individual human searchers demonstrating what the dog adds to the team's search efficacy (Chiacchia KB *et al.*, 2015).

Locations for tasks were chosen arbitrarily, following the team's training program and goals. Meteorological measures of insolation and wind were taken in the field by each handler using the methods previously described (Chiacchia KB *et al.*, 2015). Vegetation density was derived by scoring subjective descriptions of each area taken at the time of the task, using terms such as "open" or "sparse" as denoting open vegetation, "heavy" or "thick" as denoting heavy vegetation, and "moderate" or combinations of open and heavy to describe parts of the same area as moderate/mixed vegetation. Ecoregions for each task were scored by registering a map of level III and IV ecoregions in U.S. EPA Region 3 against the task locations using SARTopo (See Figure 3). Humidity and temperature data for each task were determined with the online Weather Underground site, which collects data for the nearest weather station to a given location (Online at https://www.wunderground.com/).

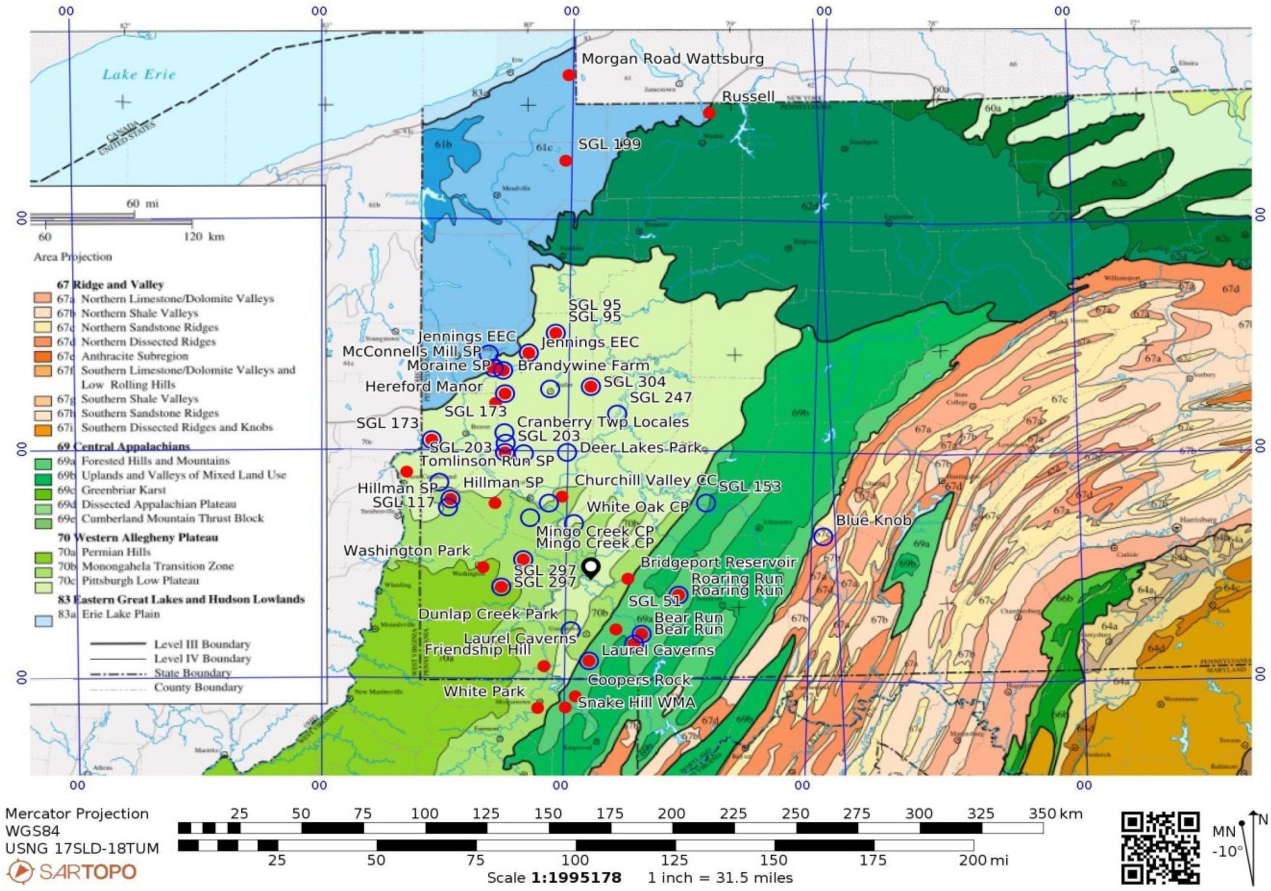


Figure 3: Location of air-scent-dog-team tasks, on a U.S. EPA level III and IV ecoregion map. Sites in the current study are solid red circles; sites in the previous study (Chiacchia KB et al., 2015) are open blue circles.

The authors collected data from 12 air-scent dog teams consisting of 2 handlers (denoted handler 1 and 2) alternating in handling 6 dogs. Note that both handlers began the study with nearly identical training and operational histories, having both become operational in December, 1992. Dog/handler pairs were determined at each training with a rotation, in an attempt to ensure that each pair was maximally represented. Due to staggered entry of dogs into the study and the occasional absence for medical reasons, each team was nevertheless not equally represented. Data collection for the study was also suspended from May through November 2019 due to an author/handler's shoulder injury and subsequent recovery. The dogs were:

- Dog A, Pipistrelle, an intact female English shepherd, born in 1999. Pip died in 2012 and so did not participate in the new data collection, though data from the previous study used in this report includes her.
- Dog B, Sophia, a neutered female German shepherd dog, born in 2005. Sophie entered the current study at its beginning, and retired from SAR and the study due to age-related medical reasons in April, 2017. She was the only dog represented in both data sets.
- Dog C, Brier Rose, a female English shepherd born in 2007. Note that Rosie is Pip's daughter and was neutered in 2019. Rosie entered the current study at its beginning.

- Dog D, Cole, a neutered male English shepherd, born in 2008. Cole entered the current study at its beginning. Unlike the other dogs in this study, which were purpose-bred for either police, farm, or SAR work, as a puppy Cole was a rescue from a felony animal cruelty case.
- Dog E, Charlotte, an intact female English shepherd born in 2013. Charlie is Rosie's daughter, and entered the current study at its beginning.
- Dog F, Verity, an intact female English shepherd born in 2017. V is Charlie's daughter, and entered the current study in December, 2019.

Dog teams, consisting of the dogs above paired with the two individual handlers, were certified via an external testing organization, the Pennsylvania Search and Rescue Council, using their 2010 air-scent testing standard (corresponding to NIMS type III non-discriminating air-scent team; online at https://img1.wsimg.com/blobby/go/8cfbf535-53eb-493c-8c8c-c88150fa0928/downloads/1d2vbidlq_354974.pdf?ver=1611603462319). The exceptions were teams 1E, 2D, and 2F (*i.e.*, handler 1/dog E, etc.), which were not certified during the duration of the current study. The certified dog-and-handler teams also worked together on multiple real SAR missions.

Human Visual Experiments

We conducted the human visual sweep width experiments in the manner of Koester RJ *et al.*, 2004, with modifications previously described (Chiacchia KB, Houlahan HE, 2010), using the IDEA Microsoft Excel worksheet provided by R. Koester and N. Guerra to automatically generate a randomized plan for a sweep-width course. We calibrated the courses either by measuring average maximum detection range (AMDR) values on site, or from previous experiments in the same area and season.

In addition to data from the two effective-sweep-width experiments from our previous report (Chiacchia KB, Houlahan HE, 2010), we used data from two additional experiments. The results from these contributed to another previous report (Koester RJ *et al.*, 2014), but have not previously been published in toto. These experiments took place at Clarence Schock Memorial Park, Mount Gretna, Pennsylvania, U.S.A., in the spring of 2006 and State Game Lands 203 near Wexford, Pa., U.S.A., in the summer of 2010.

The area chosen for the Mount Gretna experiment was the north slope of Governor Dick Hill in Clarence Schock Memorial Park. Starting point for the course was 18TUK7658057276 (U.S. National Grid, Datum NAD83); total length was 3,350m, with the last clue placed at 3,320m. The area (U.S. EPA Ecoregion 64, Northern Piedmont) is largely second-growth forest, dominated by a birch, beech, and maple canopy common to the Allegheny Mountain ridges, with a mid-story of birch saplings and spicebush and a ground cover of light bramble, birch, and spicebush saplings. The vegetation had just begun to leaf, producing perhaps a transitional situation compared with winter or summer sight lines. Another major feature of the area was frequent outcroppings of large (one- to two-meter or larger diameter) boulders of a green-gray hue quite similar to that of the green adult mannequin (see Discussion below).

The days of the experiment (May 4-6, 2006, in coordination with the Pennsylvania Search and Rescue Council's then-annual SAR-EX conference) proved largely clear, with cool mornings and mildly warm days, and little cloud cover. 26 searchers followed a course consisting of 19 real and 3 virtual high-visibility human targets, and 17 real and 2 virtual low-visibility human targets. ("Virtual" objects occurred when the random placement would have made the object completely invisible from the course. In these cases, the virtual location was counted as an automatic miss, with the real object moved toward the course until just visible to a cued experimenter.) With discounted DOs, this made for a total of 553 DOs on the high-visibility target and 485 on the low-visibility target.

The first three searchers completed the course in the late afternoon/early evening of May 4. Sundown occurred while these searchers were still on the course, making the relevance of subsequent detection opportunities (DOs) to daytime visual search questionable. Therefore, we used local sundown that day (20:03 Eastern Daylight Time) as an endpoint, counting neither detections nor misses recorded after that time. Upon our confirmation run of the course at the end of the experiment, we found one of the high-visibility mannequins to have moved onto the trail, presumably from wind. We discounted any detections or misses on this search object after the last searcher made a positive identification on it in its proper place, which had happened earlier that day. Another issue arose on the morning of May 6, when a park staff member removed one of the high-visibility mannequins near the trail (lateral distance 1m) and subsequent navigational flags. Upon clearing up the matter with the staff member (who had mistakenly thought the mannequins and flags to be unauthorized), we replaced the navigational flags and the dummy. No searchers missed this target, so apparently none had passed by its position in the time between the ranger removing it and our replacing it.

Starting point for the Wexford experiment at State Game Lands 203 was 17TNE7396998772; total length was 2,478m, with the last clue placed at 2,464m. The area (Ecoregion 70, Western Allegheny Plateau) is largely second- and third- growth forest that had been utilized for farming as well as strip mining, gas and oil drilling, and more recently has been managed for game-bird habitat via clear-cutting and resulting succession. The woods are dominated by a mixed maple, tulip poplar, oak, cherry, and hickory canopy common to the Allegheny Plateau, with a mid-story of spicebush and sassafras and a ground cover of ferns, mayapple, and naturalized multiflora rose. Occasional stands of naturalized black raspberry bramble and fox grapevines also dot the area. The vegetation was in full leaf for the experiment, with full summer growth of the brambles. A few leaves of the spice-bush plants had gone yellow, presumably due to stress from a dry spell, but the great majority of the vegetation was still green. General vegetation types included open woods; fields and mowed trail areas; brambles; and sucker stands, the latter being the five- to 10-year succession vegetation after clear-cutting.

The weather for the single day of the experiment (Sept. 11, 2010), was mild, beginning in the low 10s°C in the morning, hitting a high in the low 20s around 15:30, and staying there until sundown. Skies were clear for the duration of the experiment. 24 searchers followed the course, which was laid out intentionally not following extant trails, with search objects consisting of 17 real and 2 virtual low-visibility-mannequins; 16 real and 5 virtual orange gloves, 15 real and 3 virtual blue gloves; and 15 real and no virtual brown gloves.

When confirming the search object placements after the experiment, we found that one orange glove that had randomized to a crossing of our course with an extant trail had been moved, presumably by a passerby assuming it had been accidentally dropped and intending to place it where it would be easier to see. We discounted all detections and misses on this glove. We also discarded all data from one searcher (number 11) from whom we did not obtain a signed release form. The end result was 435 low-visibility-mannequin, 483 orange-glove, 413 blue-glove, and 345 brown-glove DOs.

The low- and high-visibility mannequins used in both of the new experiments were the same ones used in our previous report (Chiacchia KB, Houlahan HE, 2010). We obtained brown work gloves (Tractor Supply, Brentwood, TN, USA) as low-visibility clues; blue jersey work gloves (Grand Rapids Industrial Products, Grand Rapids, MI, USA), intended to be medium-visibility clues; and orange gloves (Boss Manufacturing Company, Kewanee, IL, USA), as high-visibility clues. Note, however, that the blue of the gloves was quite bright and returned a *W* value similar to that of the orange glove (see below). For the SGL 203 but not the Mount Gretna experiment, as previously reported (Chiacchia KB, Houlahan HE, 2010), we forced the IDEA worksheet to place the low-visibility objects (green mannequins and brown gloves) to a maximum distance of 1 X AMDR rather than the usual 1.5 X AMDR to avoid placing these objects too far from the course to obtain crossover plots (Koester RJ *et al.*, 2014). We also made use of the post-experiment confirmation box in the search object placement worksheet to jot down a short subjective description of the vegetation surrounding each search object.

Deriving Effective Sweep Width

In addition to determining individual teams' crossover *W* values as previously described (Chiacchia KB, Houlahan HE, 2010; Chiacchia KB *et al.*, 2015), we determined PODs for given distances in the following manner. In the random-search model:

2. $\text{POD} = 1 - e^{-C}$

where C, or coverage, is defined as:

3. $C = Wd/A$

W being effective sweep width for a particular detector (or group of detectors) and search object under given conditions, d the length of the detector's path within the search segment, and A the area of that segment.

Each lateral approach of a detector past a search object can be considered as an area defined by:

4. $A = 2Ld$

where L is the lateral distance at the detector's closest approach to the search object, and again d is the length of the detector's track. We use 2L because *W* is measured from the detector's right and left as it moves through the search area. When a DO occurs, then, its lateral distance thus defines a coverage for that DO and lateral distance. A POD can be determined for each value of C, then, by cumulating any detections between the detector's path and that value of L and dividing them by the number of detections divided by detections plus misses within that L (See Figure 4, modified from Cooper DC *et al.*, 2003).

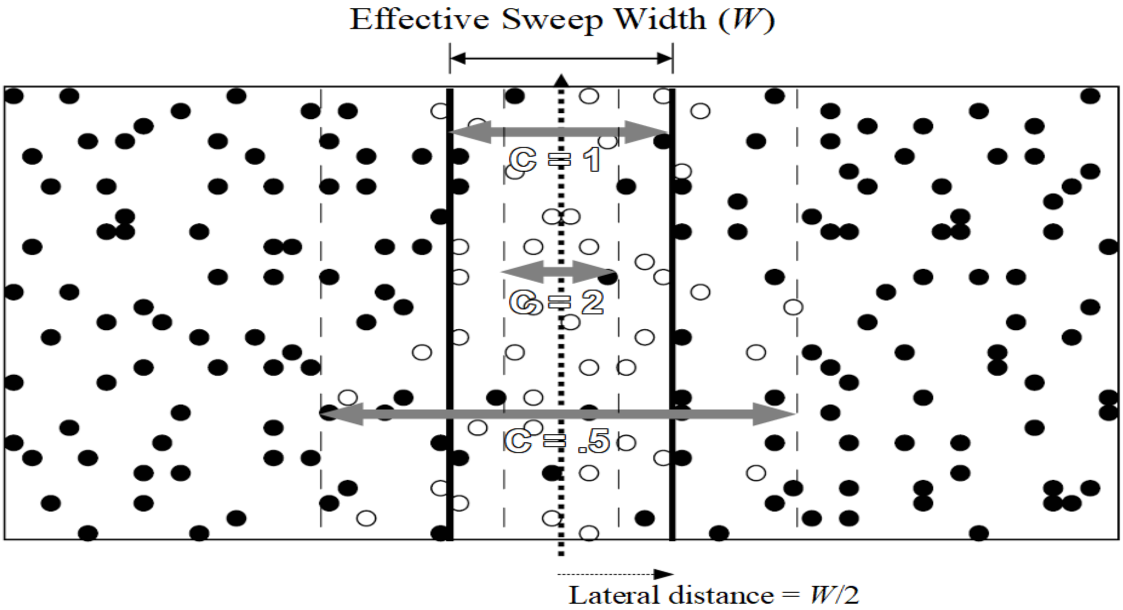


Figure 4: Determining coverage for individual detection opportunities (DOs) using the envelope defined by lateral distance (L). Black circles represent misses, open circles detections. Gray arrows and dashed lines show the envelope that would define coverages of 0.5, 1, and 2, each defined as 2L/W; the dotted line/arrow shows the detector's path.

For each pass, and substituting equation 4 into equation 3 and then 3 into 2, the POD equation, then, becomes:

5. $\text{POD} = 1 - e^{-(W/2\text{L})}$

We want to derive *W*, which we do not yet have. Since the coverage for a detector with an arbitrary *W* of 50m at a lateral distance L would be:

6. $C_{50} = 50\text{m}/2\text{L}$

for a detector of unknown *W* we can then substitute:

7. $\text{POD} = 1 - e^{-[(50\text{m}/50\text{m})(W/2\text{L})]}$

to get:

8. $\text{POD} = 1 - e^{-(C_{50} W/50\text{m})}$

Thus, by plotting POD for a given detector and group of DOs against C_{50} — again, the coverage at the distance of a given DO assuming a *W* of 50m — we can derive the detector's actual *W* as a constant from a least-square fit of the data against equation 8.

W may be derived in the same way using different models for the relationship between POD and *W*, such as the "inverse-cube" model (Koopman BO, 1980):

9. POD = erf[(√π)C/2]

where erf is the statistical error function and C is coverage as above (also see Results and Discussion).

The question of how far from the detector's path to begin counting DOs (*i.e.*, how high a C_{50} to include in the curve) is important, as the small numbers of DOs close to the detector would result in high noise of POD due to low N (the noise in Figure 2 stems from the same source), and thus the points in this part of the graph will have a high degree of variance. At low C_{50} we have a different problem: Because of the finite size of the search areas in the study, longer Ls will become sparse and, because we are cumulating the POD from the detector's path outward, if there is a gap between two DOs/Ls we will have the same decrease in measured POD as if the DOs were close together. We would thus expect the data points to artifactually begin to "lift off" of the curve at low C_{50}. To partially ameliorate both problems, we selected an arbitrary number of DOs both from the detector outward and from the longest DO inward before beginning to count and tally detections and misses for the POD determination.

The issue is how best to select that arbitrary number. In the current study, we leveraged the limitations of the human visual sweep width data. Our second such experiment, in summer 2006, involved 18 human searchers, the smallest number in our visual-search datasets (Chiacchia KB, Houlahan HE, 2010). Because DOs in the human experiments are number of searchers multiplied by number of search objects, setting the minimum number of DOs below 18 would have meant discarding results from two each of the 23-29 search objects of each type placed in that experiment — a major loss of data. Thus, we set the minimum number of DOs to begin calculating POD both from the detector outward and the longest DO inward at 18 for both canine and human data.

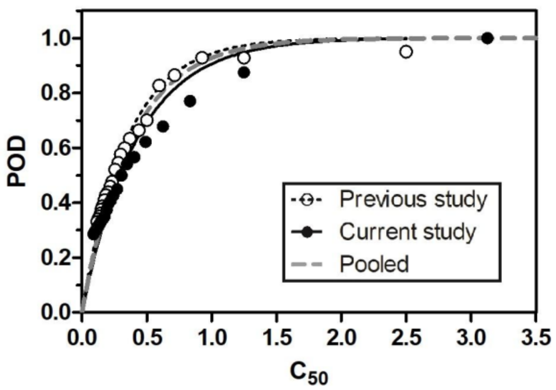
Statistical Methods

The current report employed GraphPad Prism version 5.00 for Windows (GraphPad Software, San Diego, CA, USA, www.graphpad.com) to perform statistical testing. Data are expressed as mean±standard deviation (SD). All tests were two-tailed, with $P<.05$ set as the threshold for statistical significance. We used Microsoft Excel and Apache OpenOffice to create a worksheet that performed simple calculations, including correcting P-values for repeated testing using the step-down Hochberg/Šidác method.

Results

Air-scent Dog Teams

Of the 167 included tasks in the new data, 10 resulted in no find before intentional unblinding, and a further 7 resulted in refind failure (*i.e.*, the dog found but did not tell the handler and so it was counted as a miss). The resulting PODs are plotted against C_{50} in Figure 5a. We also plotted the data from our previous report (Chiacchia KB *et al.*, 2015). For clarity, Figure 5a only shows one point out of every six, though all points were included in this and subsequent analyses.



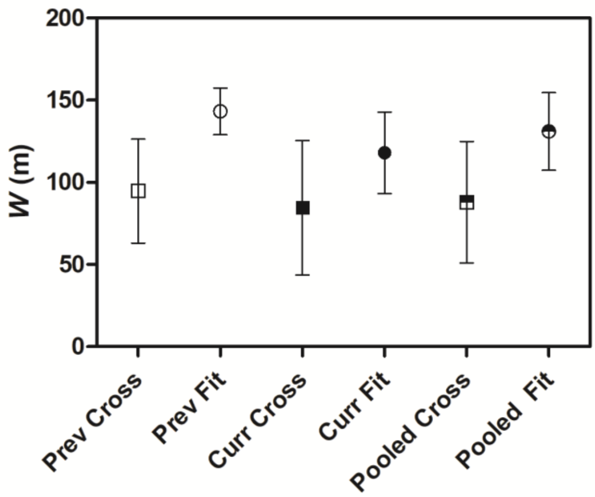


Figure 5a (left): Plots of C_{50} (coverage assuming W *of 50m) versus probability of detection (POD) and the resulting least-square fits of the air-scent data to the random search model. Data shown are from the previous study (Chiacchia KB et al., 2015), open circles and dotted line; the current study, black circles and solid line; and the fit to both sets of points, thick gray dashed line. For clarity, only every sixth point is pictured.*

5b (right): W *values for the previous, new, and combined data. Open symbols are the previous study; black symbols current; and half-black, half open combined data. Squares are* W *determined by the crossover method; circles are from the fits in 5a. Shown are* W±*SD.*

The result of the least-square plot is that *W* for the new data is 118±25m, with an R-square of .895 and an N, after excluding points at the upper and lower part of the graph as described in Methods, of 138. These 138 points in the graph are underlain by 526 included DOs (with the POD at the lowest C_{50} value calculated from the full 526 DOs, and the highest from 18 DOs). The data from the previous report returned a *W* of 143±14m, R-square .978, N=139, and 320 DOs. An extra sum-of-squares F-test produced P<.0001 (P-value corrected for repeated testing <.005) for the null hypothesis that the resulting curves share a common *W* value, despite the very small difference in POD between the two curves at all points (on the order of ≤~5%).

One factor to be kept in mind is that the 50% POD level may be a "tipping point" for shifting search efforts to new areas, when using shifting POAs (see Introduction and Charnes A, Cooper WW, 1958). Thus it is useful to observe the difference in C_{50} at the curves' 50% POD level: as $C=Wd/A$ (Equation 3) and $d=vt$ (the detector's velocity multiplied by time searching in a given area), differences in C_{50} in these curves will be proportional to the time needed to reach a given level of POD.

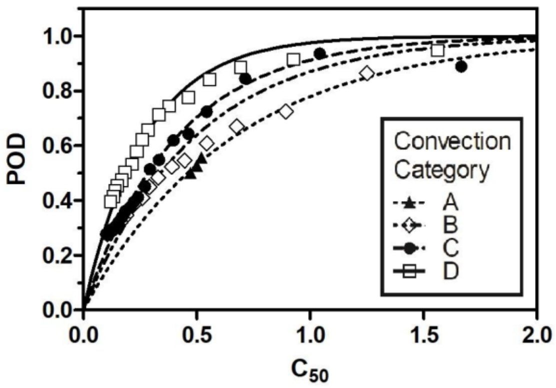
A Wolfowitz runs test for deviation from the random-search model was significant, with $P<.0001$ (corrected to $<.004$ for the current data, $<.005$ for the previous curves; see Discussion for more on this). Because of the small POD difference between the two datasets, we also did a fit of the two sets of points pooled together. The result was W=131±24m, R-square .926, N=268, and 846 DOs.

Figure 5b shows the resulting "fit" W values compared with those acquired from averaging crossover plots of individual dog/handler teams for the current, previous, and pooled data. The latter were, respectively, 85±41m, 95±32m, and 88±40m, with Ns of 8 (because not all 10 teams in this group had sufficient data for a successful crossover plot), 4, and 12 (because for the purpose of this last comparison dog B's data were combined, instead of divided between the previous and current studies). These values were smaller than those derived from the least-square fits, and the uncertainties of the mean W values were higher, phenomena examined below.

The unlikelihood that a POD difference of ~5% will ever be operationally significant (See Discussion) and the excellent R-square value of the pooled POD data against C_{50} (and thus W) led the authors to pool the data between the two datasets in subsequent comparisons, to achieve higher statistical power.

Atmospheric convection proved a significant correlate with W in the authors' previous report (Chiacchia KB *et al.*, 2015; the rationale being that convection causes turbulence that dilutes scent, see Graham H, 1994). Therefore, we addressed this issue again, plotting POD for the pooled data against C_{50} in the different daytime convection conditions (Figure 6a) — A through D, with A representing the strongest convection and most hypothesized disruption of scent transport (*i.e.*, lower W). Again, for clarity, only every sixth point is shown for conditions B through D in Figure 6a, though all were included in the analysis. All points are shown for condition A.

In Figure 6a and 6b, we see a correlation of stronger convection with poorer search performance. Ws for convection conditions A, B, C, and D were, respectively, 75±3m, 102±21m, 121±12m, and 184±18m, with R-squares of .856, .908, .980, and .971. Note that the low SD for the A value is likely an artifact of the scant data for that curve; Ns for each condition were, respectively, 4, 127, 119, and 100, with 20, 353, 253, and 204 DOs. An ANOVA of the parameters generated by the fit returned $P<.0001$ (corrected $<.002$); a Tukey's multiple comparison post-test showed $P<.0001$ for differences between each convection condition except A vs. B, which was $P<.05$. The fit for convection conditions B through D displayed a significant runs test ($P<.0001$, corrected to $<.004$ for each) but the fit for A did not ($P=.667$, corrected .963).



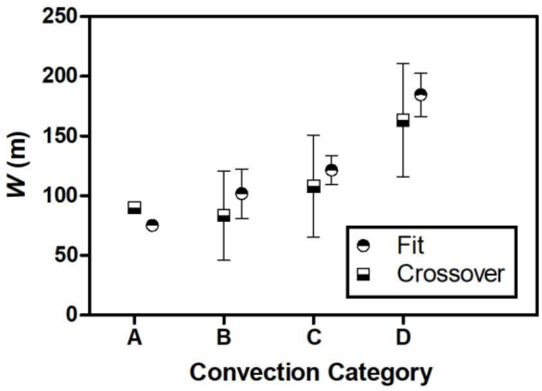


Figure 6a (left): Plots of C_{50} versus POD for the pooled air-scent data at the four measured convection conditions (A, strongest convection, D, weakest). For clarity, only every sixth point is pictured except for A, in which all points are shown.

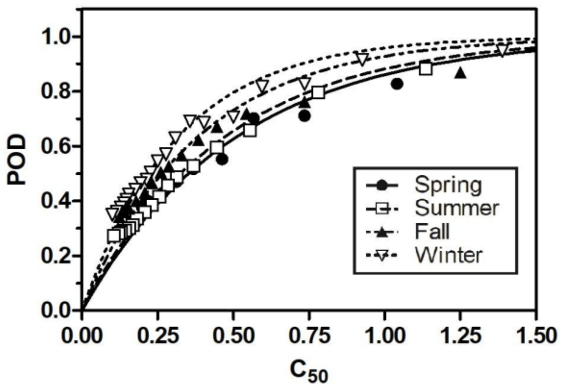
6b (right): W *values at the different convection conditions. Squares are W determined by the crossover method; circles are from the fits in 6a. Shown are* W±*SD.*

Also shown in Figure 6b are the pooled data from these comparisons as calculated via crossover plots. Again, the crossover *W*s were both smaller and had larger errors than the fit values: respectively, 90m, 83±37m, 108±43m, and 163±47m. Ns were 1, 10, 7, and 9, respectively. The A-convection crossover value has no SD because only one of the teams possessed enough data in these conditions to generate a crossover plot; for its fit value, the SD was smaller than the size of the circular symbol. Because of the additional data, the crossover values in this figure differ slightly from those in the otherwise equivalent Figure 6 in our earlier report (Chiacchia KB *et al.*, 2015).

Human visual-searcher *W* values correlate strongly with season (Chiacchia KB, Houlahan HE, 2010). Thus, any relationship between air-scent dog performance and season would be useful for comparing the two search modalities. The authors' previous report did not detect such a relationship for air-scent dog teams (Chiacchia KB *et al.*, 2015), but the smaller SDs of the fit *W* values and the increased Ns of the pooled data encouraged us to make the comparison with the expanded data set.

The result of that comparison can be seen in Figures 7a and b. Again, in 7a only every sixth point is shown. The resulting *W*s for summer, fall, winter, and spring were 106±15m, 131±19m, 158±21m, and 98±11m;

R-square .956, .912, .949, and .961; N=105, 84, 118, and 59; and 236, 172, 225, and 143 DOs. A Tukey's post-test of an ANOVA of the parameters (P<.0001, corrected <.003) returned P<.001 for all pairwise comparisons except for spring versus summer, which was P<.05. The difference between the high and low curves (winter versus spring) approached the ~20% level at C_{50} values of approximately .5. (Note this would relate to an actual C of around 1 or higher using the specific *W* values in each season.) Runs tests were significant for summer and fall, P<.0001 corrected to <.004 for each, but spring lost its significance upon repeated-measures correction (P=.003, corrected .067); neither was the winter runs test significant (P=.361, corrected .932).



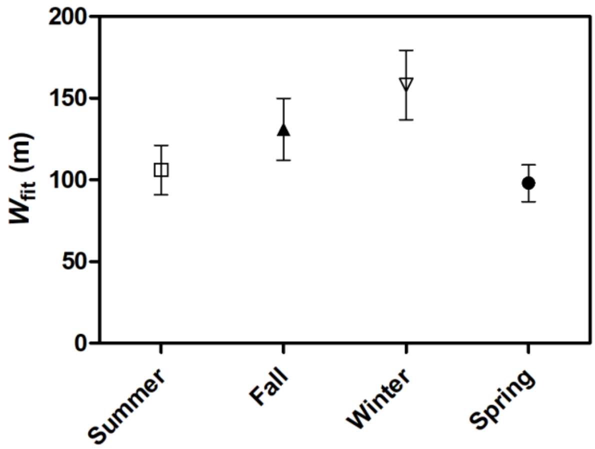
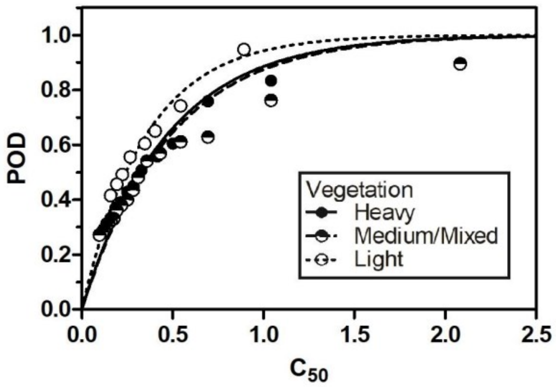


Figure 7a (left): Plots of C_{50} versus POD for the pooled air-scent data in the different seasons. For clarity, only every sixth point is pictured.

7b (right): W *values in the different seasons from the fits in 7a. Shown are* W±*SD.*

The authors previously detected a difference among *W*s in human visual search for search objects randomized to areas subjectively judged as heavy versus light vegetation (Chiacchia KB, Houlahan HE, 2010). For another possible correlate of air-scent performance that allows apples-to-apples comparison with human searchers, we investigated the relationship between POD and C_{50} (and thus *W*) versus subjective vegetation thickness (see Methods), shown in Figures 8a and b (again, only every sixth point shown for 8a). As this information was not collected in our earlier study, this comparison relied only on the new data. For heavy, medium/mixed, and light vegetation, the resulting *W*s were respectively 108±13m, 103±23m, and 142±17m; R-square .939, .866, .918; N=67, 107, and 45; and 127, 78, and 96 DOs. In a Tukey post-test of an ANOVA (P<.0001, corrected <.003), the light-vegetation *W* was significantly larger than each of the others (P<.001), but the heavy and medium/mixed values were not significantly different. The runs tests lost their significance upon correction (heavy vegetation P=.002 corrected .052; medium/mixed P=.005 corrected .088; light P=.031 corrected .331).



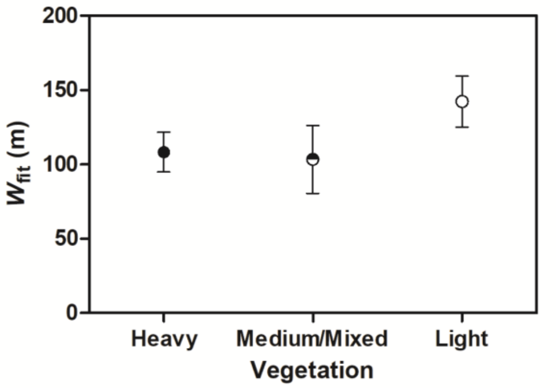
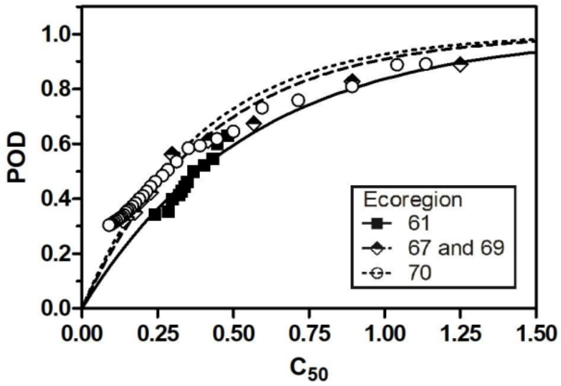


Figure 8a (left): Plots of C_{50} versus POD for the pooled air-scent data in different ground-level vegetation thickness scored subjectively (heavy, light, and either medium or a mix of heavy and light). For clarity, only every sixth point is pictured.

8b (right): W *values in different vegetation thickness from the fits in 8a. Shown are* W*±SD.*

Another variable that seems likely to be relevant to air-scent dog team performance is the U.S. EPA ecoregion in which the task takes place. Ecoregions differ by climate, elevation, vegetation, and terrain in ways that will likely impact both scent transport and handlers' choices of search tactics. Thus we compared POD curves and resulting *W*s for the ecoregions in which our study tasks were undertaken — which again were arbitrarily chosen and not equally distributed. The largest such number of tasks (199) was in Ecoregion 70, the Western Allegheny Plateau. For the next region, we pooled 67 tasks in Ecoregions 67 and 69, respectively Ridge and Valley (specifically 67c, Northern Sandstone Ridges) and Central Appalachians, which are subjectively similar rugged-terrain regions. Finally, a smaller number of tasks (13) took place in Ecoregion 61, Erie/Ontario Drift and Lake Plains. (See Figure 3 for the map of tasks and ecoregions.)

In Figures 9a and b we see the results (every sixth point shown in 9a except for Ecoregion 61, in which all are shown), with *W*s for Ecoregions 70, 67/69, and 61, respectively, of 130±26m, 120±17m, 90±8m; R-square .923, .942, .904; N 199, 67, and 13; and 677, 123, and 35 DOs. A Tukey post-test of an ANOVA (P<.0001, corrected <.003) gave P<.001 for 70 versus 61 and 67/69 versus 61 and P<.01 for 70 versus 67/69. Runs test P-values were <.0001 corrected <.003 for Ecoregion 70 and Ecoregions 67/69, and .003 corrected .054 for Ecoregion 61.



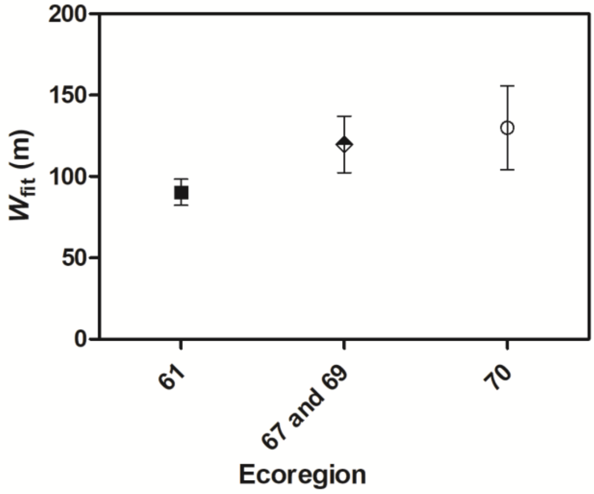


Figure 9a (left): Plots of C_{50} versus POD for the pooled air-scent data in the different U.S. EPA ecoregions. For clarity, only every sixth point is pictured for Ecoregions 70 and 67-and-69; all points are shown for Ecoregion 61.

9b (right): W *values in the Ecoregions from the fits in 9a. Shown are* W*±SD.*

Another variable of interest in the dog-handler community is that of humidity, which arguably could increase scent via bacterial metabolism creating more odorants. As the authors did not collect humidity data in the field, this comparison required the use of historical weather data. As the authors also did not in their initial report retain start and end times for each task once sun angles/convection were calculated, this comparison also was limited to the new data. We separated the data based on three humidity ranges of as similar number of DOs as possible, to obtain the best separation between the different conditions while maximizing the number of DOs for each.

Figures 10a and b show a comparison of the POD curves and the resulting *W* values, respectively. For the 20-49% low-humidity range, *W* was 114±20m; mid-humidity >49-67%, 106±21m; and high-humidity 68-100%, 113±16m. An ANOVA of the difference between these values returned P=.030, corrected .346 (R-square for each fit .905, .851, and .928, respectively; N 72, 68, and 81; DOs 156, 157, and 162). Runs test P-values were <.0001, corrected to <.004 for each curve.

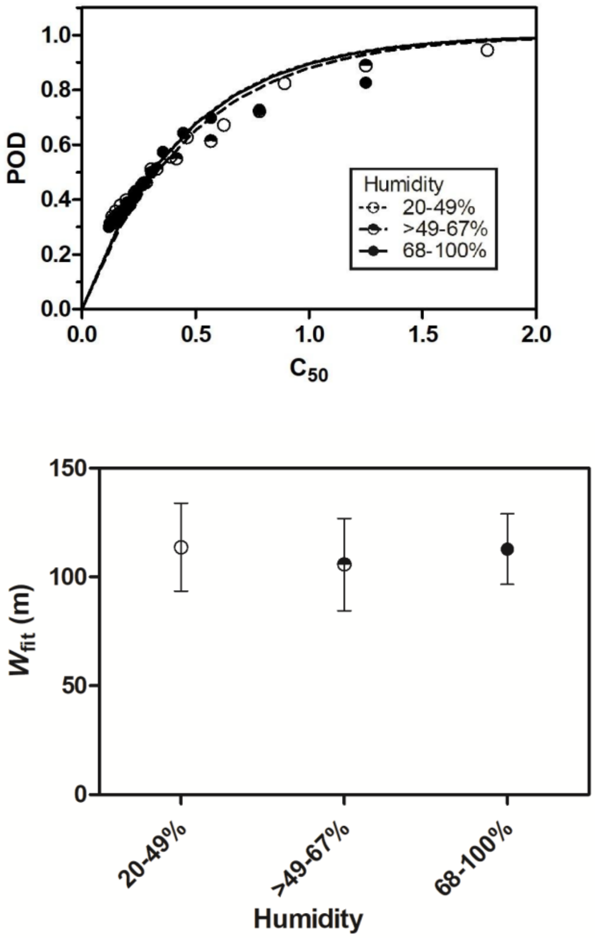
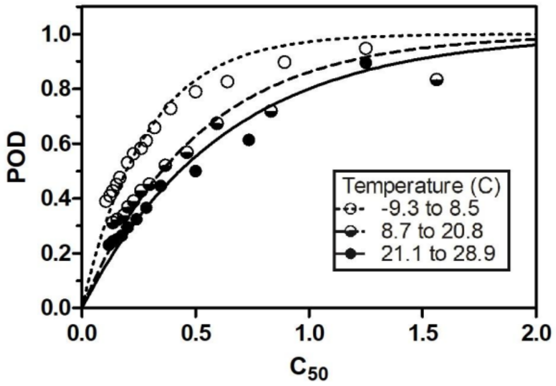


Figure 10a (left): Plots of C_{50} versus POD for the pooled air-scent data at different levels of humidity. For clarity, only every sixth point is pictured.

10b (right): W *values at the humidity levels from the fits in 10a. Shown are* W±*SD.*

Temperature is a matter of great concern to dog handlers, dogs being much more vulnerable to heat illness than humans. Subjectively, handlers note that dogs are less comfortable and may work less effectively at high temperatures. For that reason, temperature is another promising factor for determining *W* for dog teams. As with humidity, temperatures were determined from historical records as described in Methods. The authors compared POD curves and resulting *W* values from the fits in Figures 11a and 11b, respectively, again dividing the tasks into 3 temperature spans of as roughly equal size as possible in DOs (once again only every fifth point pictured in 11a).



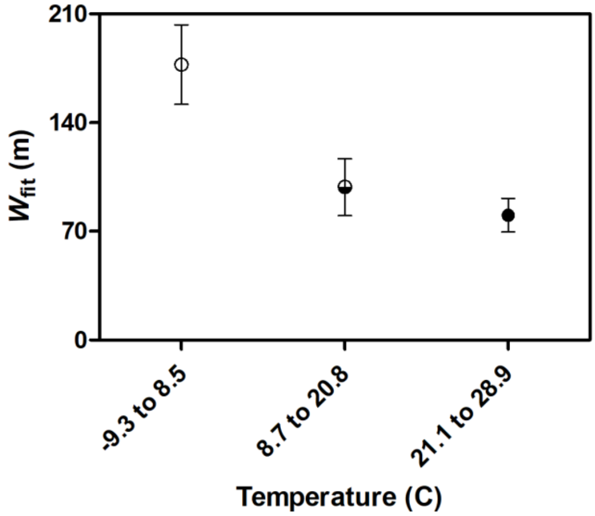


Figure 11a (left): Plots of C_{50} versus POD for the pooled air-scent data at different temperature ranges (C). For clarity, only every sixth point is pictured.

11b (right): W *values at the temperature ranges from the fits in 11a. Shown are* W±*SD.*

As can be seen in Figure 11a, the POD separation between the highest-POD (lowest T) and lowest-POD (highest T) curves is fairly large, on the order of the separation seen for convection in Fig. 4a. For the lowest temperatures, -9.3 to 8.5ºC, *W* was 177±26m (R-square .930; N 86; DOs 159). For the medium range, 8.7 to 20.8ºC, *W* was 99±18m (R-square .868; N 70; DOs 153). For the highest temperatures, 21.1 to 28.9ºC, *W* was 80±11m (R-square .955; N 62; DOs 169). A Tukey post-test of an ANOVA (P<.0001, corrected <.003) revealed significance comparing all of the temperature ranges against each other (P<.001). Runs test for the low-temperature curve was P<.0001 corrected to <.004, the others <.0001 corrected <.003.

Human Visual Searchers and Combined Analysis

The authors applied the curve-fit method for deriving *W* both to data from our previous report for visual human searchers at a single location on the Western Allegheny Plateau (Chiacchia KB, Houlahan HE, 2010) in the summer and winter as well as new data collected at that same location in the summer and a location in the spring in Ecoregion 64, Northern Piedmont, in Mount Gretna, Pennsylvania.

Figures 12a through c plots the POD versus C_{50} data and fits for the human experiments. Figure 12d compares the *W*s derived from these curve fits versus crossover method. *W,* SD, R-square, and runs-test P-values and corrected P-values for each can also be seen in Table 1 (the table also provides the R-squares for an inverse-cube-model fit to the data, see below). Note that for two of the fits, P-value corrections removed the significance of the runs test (the orange glove in the SGL 203 experiment P=.025 corrected to .316; the low-vis mannequin in the Mount Gretna experiment P=.012 corrected to .195); the others were not significant before correction.

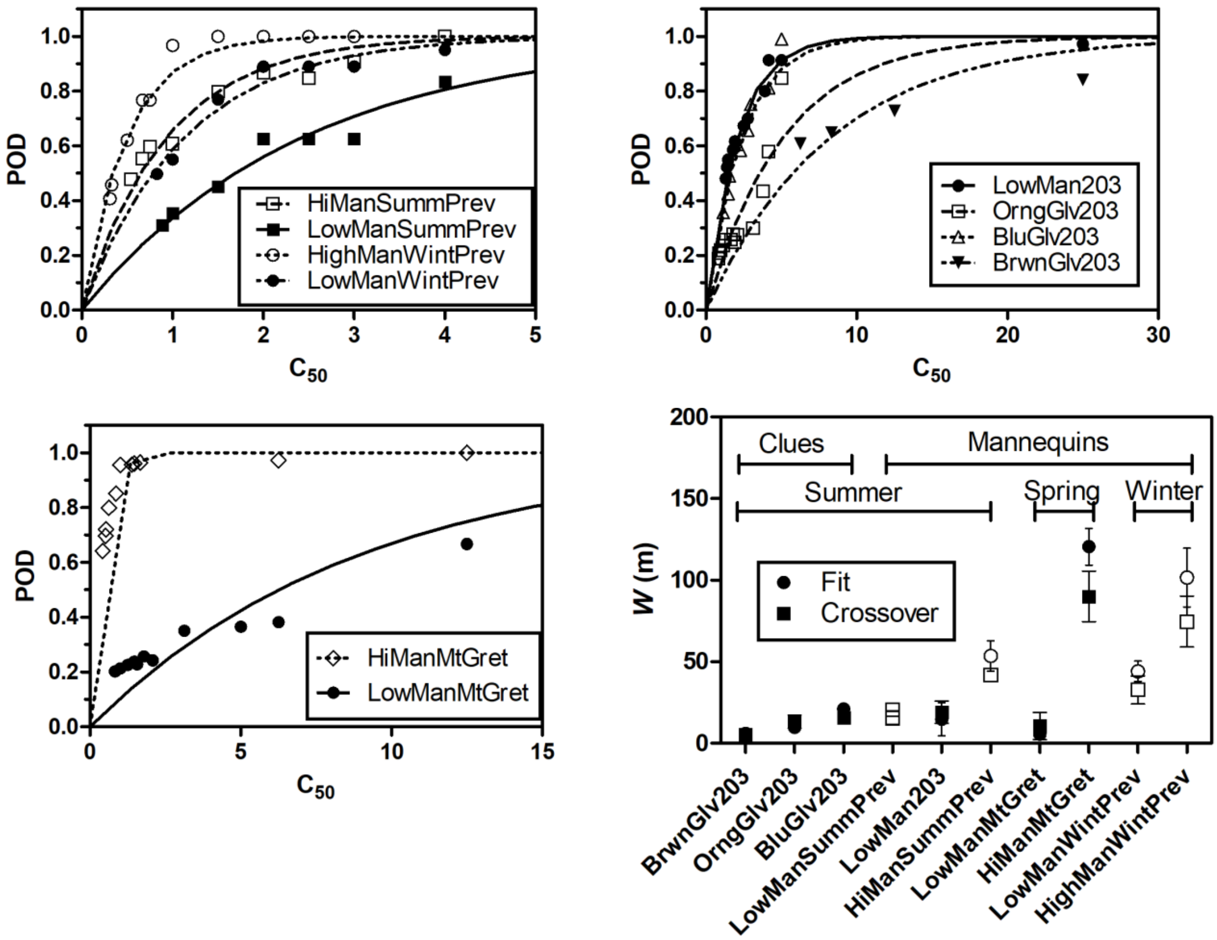


Figure 12a (top left): Plots of C_{50} versus POD for the previously reported human visual search values. HiManSummPrev: the high-visibility (white and blaze orange) mannequin in the previous summer State Game Lands 203 experiment; LowManSummPrev: the low-visibility (olive drab) mannequin in the summer experiment; HighManWintPrev: the high-vis mannequin at the winter 203 experiment; LowManWintPrev: the low-vis mannequin in the winter experiment.

12b:(top right): Plots of C_{50} versus POD for the newly reported summer SGL 203 experiment. LowMan203: the low-vis mannequin in the new 203 experiment; OrngGlv203: the orange glove at that experiment; BluGlv203: the blue glove at that experiment; BrwnGlv203: the brown glove at that experiment.

12c (bottom left): Plots of C_{50} versus POD for the spring Mount Gretna experiment. HiManMtGret: the high-vis mannequin at that experiment; LowManMtGret: the low-vis mannequin at that experiment.

12d (bottom right): W values for detection of different search objects by human visual searchers at different sites, in different seasons. Values from previous studies are open symbols, the current study black; crossover-method values are squares, random-search-fit values circles.

Table 1: Fit (random-search model) versus crossover W *values for human visual searchers and R-square for inverse-cube-model fits (*W *and SD in meters).*

	Low-Vis Mannequin Summer SGL 203			Orange Glove Summer SGL 203			Blue Glove Summer SGL 203			Brown Glove Summer SGL 203			High-Vis Mannequin Spring Mt. Gretna			Low-Vis Mannequin Spring Mt. Gretna			High-Vis Mannequin Summer Previous Study			Low-Vis Mannequin Summer Previous Study			High-Vis Mannequin Winter Previous Study			Low-Vis Mannequin Winter Previous Study		
	W	SD	N	*W*	SD	N	*W*	SD	N	*W*	SD	N	*W*	SD	N	*W*	SD	N	*W*	SD	N	*W*	SD	N	*W*	SD	N	*W*	SD	N
Random-search fit value	14	10	14	10	4	15	21	3	12	6	3	7	120	11	12	6	2	12	54	9	10	21	3	8	101	18	11	44	6	8
Crossover value	19	7	24	13	4	24	16	3	24	5	2	24	90	15	26	11	8	26	42	4	18	15	4	18	75	16	20	33	9	20
R-square of fit	.975			.868			.966			.264			.973			.555			.933			.939			.964			.956		
P for runs test of fit (corrected)	.070 (.581)			.025 (.316)			.643 (.984)			.130 (.752)			.833 (.833)			.012 (.195)			.071 (.555)			.800 (.960)			.191 (.852)			.200 (.832)		
R-square of inverse cube fit	0.913			0.878			0.973			Not converged			0.938			0.290			0.794			0.864			0.991			0.928		

To further investigate the relationship between the fit and crossover values for *W*, we plotted the crossover versus fit *W* values for the new and previous dog data in Figure 5 as well as all the human *W*s in Figure 12. That plot can be seen in Figure 13. The apparent proportionality of the crossover and fit *W*s led us to carry out a linear least-square fit of the data points, with the Y intercept constrained to 0. We obtained a slope of 1.44±.24, R-square .904. Note, however, that while the R-square appears to suggest reasonable predictability in the relationship, the runs test showed P<.0001 (corrected <.002), indicating that the relationship only approximates linearity.

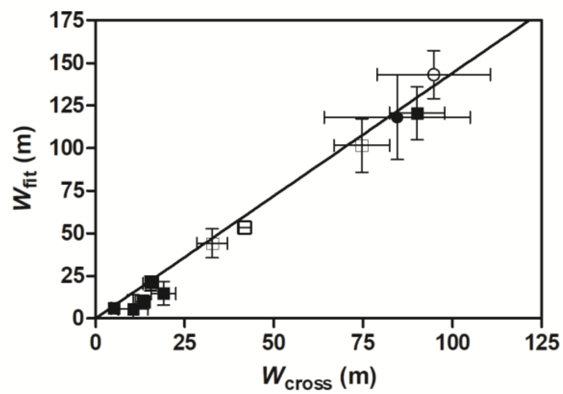


Figure 13: Plot of the crossover values of W *(*W_{cross}*) versus the random-search-fit values (*W_{fit}*) and the resulting least-square-fit line. Previous studies are open symbols; current are black. Human visual searcher values are squares, air-scent dog teams circles.*

While the SDs for the fit *W*s were smaller than those for the crossover values in the dog data, this was no longer apparent for the human data. In an attempt to clarify this issue, we conducted a paired, nonparametric Wilcoxon signed rank test of the fit versus the crossover SD values for the new and previous dog *W*s in Figure 5 as well as all the human *W*s in Figure 12, a comparison that can be seen in Figure 14. The result was not significant (P=.339, corrected .945), though again the dog SDs (circles in Figure 14) do subjectively appear to be higher for the crossover values.

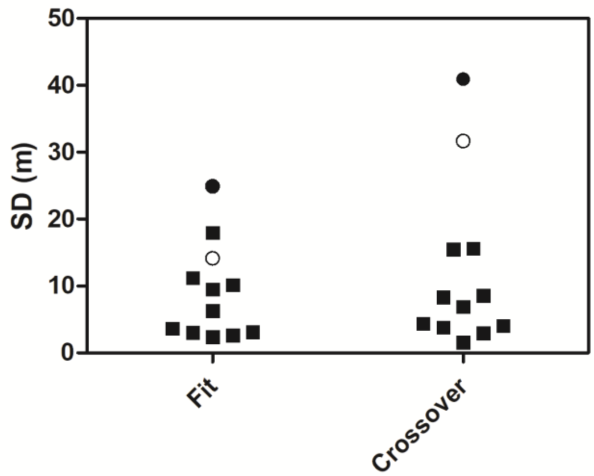


Figure 14: SD values for the random-search-fit W*s versus the crossover* W*s. Human values from all experiments are black squares; air-scent values are circles, with the previous study open and the current study black.*

Appropriateness of Random-Search Model for Ground-Based Search Efforts

The data for the air-scent dog teams as well as for the human searchers fit well to the random-search model in terms of R-square values, showing R-squares mostly >.8. Nonetheless, judging by runs tests, the model often fails (see above). In addition, fits to this model return a *W* value that is larger than the crossover value (by 1.44-fold, Figure 13 and above). This is curious given that the latter method defines *W* (Koopman BO, 1980; Cooper DC *et al.*, 2003; Koester RJ *et al.*, 2004), suggesting that the model may systematically underestimate POD and thus the fit must generate an inflated *W* to fit the data. For this reason, we compared the fit provided by the random-search model with that of another, the inverse-cube model (Koopman BO, 1980).

Figure 15 shows a plot of the new and previous dog POD data, each plotted against C_{50} and separately fit against each model. As reported above, the values for *W* using the random-search model were, respectively, 118±25m, R-square .895; and 143±14m, R-square .978. Applying the inverse-cube model, we find corresponding *W* values of 90±26m, R-square .741; and 107±19m, R-square .907. As with the random-search model, in both cases the inverse-cube model failed a runs test for deviation from the model (each P<.0001, corrected <.003). By comparison, the crossover values for each air-scent data set were 85±41m and 95±32m.

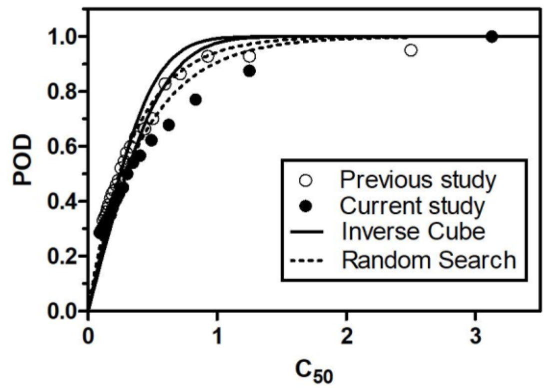


Figure 15: Plot of C_{50} versus POD for the previous (open circles) and current (black circles) air-scent data, fitting each separately to either the inverse-cube (solid line) or random-search (dotted line) models. For clarity, only every sixth point is shown.

Applying an Akaike's information criteria (AICc) test to select the model most likely to have generated the data, we find the results shown in Table 2. (The inverse-cube fit against the brown glove at the SGL 203 experiment did not converge and so that search object was not included.) The mean value for the probability that the random-search model generated the data is >75±37%, compared with the inverse-cube model.

Table 2: AICc test results for the random-search versus inverse-cube models against each dataset

Search object/experiment	*Probability of model generating data*	
	Random Search	*Inverse Cube*
Low-Vis Mannequin SGL 203	>99.9%	<.1%
Orange Glove SGL 203	35.8%	64.3%
Blue Glove SGL 203	22.1%	77.9%
Brown Glove SGL 203	n/a	Not converged
High-Vis Mannequin Mt. Gretna	98.9%	1.1%
Low-Vis Mannequin Mt. Gretna	92.9%	7.1%
High-Vis Mannequin Summer SGL 203 Previous Study	99.4%	0.6%
Low-Vis Mannequin Summer SGL 203 Previous Study	94.2%	5.8%
High-Vis Mannequin Winter SGL 203 Previous Study	0.1%	99.9%
Low-Vis Mannequin Winter SGL 203 Previous Study	84.5%	15.5%
Dog Teams Previous Study	>99.99%	<.01%
Dog Teams Current Study	>99.99%	<.01%
Mean	**>75±37%**	**<25±37%**

Discussion

The authors' intent in this report was to establish a new means of deriving *W* that offers an explicit link to field-derived POD values, smaller uncertainties, stronger statistical inference enabling comparison of more effectors of search efficacy, and the ability to distinguish between competing models for deriving POD from *W*. While our analysis accomplished most of these goals, the instances in which it did not are interesting and worthy of further study (see below).

The reader will note that, in the dog data above especially, the data points tend to sag below the fit lines at high C_{50} values. This phenomenon may indeed reflect a shortcoming of both models when applied to air-scent dog teams. On the other hand, the points here derive from a relatively small number of DOs (<~50), and thus have a noise level similar to that of the points in a lateral range curve (Figure 2). The reader should also bear in mind that there are many more points toward the low-C_{50} side of the graph, and so these low-C_{50} values (which themselves are based on high DOs and so are likely to be less affected by low-N variance) dominate the fit statistically. Another factor to be considered is that this difference may not be operationally significant as it lies in the ≤~10% range throughout the curves (see below for more on this). Ultimately, when using the shifting POA method for resource allocation, the critical point is at a POD of ~50% — that is when we would mathematically expect the method to begin strongly shifting future efforts from current high-POA areas to new areas (Charnes A, Cooper WW, 1958). Importantly, the data at this POD do not deviate widely from the curves in terms of absolute percentages in any of the fits above.

One important lesson offered by the POD curves comes from comparing the curves at the 50% POD mark, as the time required to reach that level of POD for a given detector, search object, and set of conditions is proportional to C_{50} (C being a function of d which in turn is a function of time spent searching, see Equation 3 and Results above). Thus, comparing the C_{50} of each curve allows us to see the concrete impact of that difference on time necessary to reach the arguable "tipping point" for shifting POAs.

For air-scent dog teams, the previously reported data (Chiacchia KB *et al.*, 2015) and those from the current study are in general agreement, as judged by the small (≤~5%) difference in PODs along the C_{50}/POD curve. An F-test comparing the two datasets did return a significant difference between the resulting *W*s at the chosen $P<0.05$ level. The reader can find a discussion of the difference between statistical and operational significance below; but the small POD difference convinces the authors that the operational significance in this case would likely be nil (*i.e.*, we suspect that search managers would seldom if ever find a POD difference of less than 5% to be actionable). More broadly, the POD-curve method for deriving *W* appears to provide good predictive agreement with the data and possibly smaller SD values, at least for the dog data (though see below re. the latter).

The authors used ANOVA and other tests associated with the POD-curve fits to gauge the significance of various effectors of search efficacy either reported previously (Koester RJ *et al.*, 2004; Chiacchia KB, Houlahan HE, 2010; Chiacchia KB *et al.*, 2015) or subjectively believed to be important among operational dog handlers. Of these, and for the dog teams, the three most effective comparisons (*i.e.*, providing the largest significant difference between the compared groups) were, in decreasing magnitude of difference, atmospheric convection, temperature, and season.

In a perfect world, the authors would recommend using convection to track search efficacy operationally for air-scent dog teams precisely because it offers the largest POD separation between the comparison groups. In addition, these measurements seem more likely to apply to disparate climates and terrains than, for example, the seasonal measurements. However, few operational dog handlers formally track cloud, wind, and sun angle differences through a task. While our own experience shows that such measurements are indeed possible using a minimum of extra equipment (thanks largely to the efforts of meteorologists in the wildland firefighting service — Lavdas LG, 1976; Lavdas LG, 1997) and the ubiquity of smartphone apps for determining sun-altitude angle, expanding this practice to the larger community may be a heavy lift among dog handlers already trying to formulate search tactics, read moment-by-moment wind direction, interpret dog behavior, and navigate through an unfamiliar area.

Perhaps a bigger limitation of the use of convection is it can only be done retrospectively; weather forecasts are neither accurate nor granular enough to capture the wind and cloud conditions likely to be encountered on a given search task. While dog handlers can report convection conditions on debriefing at the end of a task — information that is helpful for retrospective analysis of search efforts and thus planning of future tasks — they are not useful for prospective planning (*i.e.*, search managers telling dog handlers what level of POD they would like to achieve in a given task and thus selection of tactics to meet that level).

Because of this, temperature may be an attractive alternative to convection. Indeed, convection drives daytime air temperature (either locally or in generating the temperature of winds accompanying weather fronts) and thus the two are interrelated. Temperature data are much more simply collected. Unlike convection conditions, temperature is more stable over time and likely to be comparable between the command post and the usually close-by area of a given task, making it unnecessary for field teams to collect the data. In addition, temperature implicitly collects another factor subjectively believed by handlers to be important: The dog's comfort level, as dogs are clearly more animated and comfortable when the temperature is not high. However, to some extent the inaccuracy of weather forecasts (particularly with respect to arrival of fronts that change the temperature quickly) limit temperature as a prospective planning tool just as it does with convection.

For prospective planning, the seasonal differences are thus likely to be the most useful among those reported here. They provide an easily derived value for *W* that search planners and dog handlers alike can use to make prospective decisions (albeit possibly only relevant to the U.S. Northeast/Appalachian region in which they were derived and areas climatically similar — see the limitations section below).

Just as temperature and convection are tightly related, so are convection, temperature, and season. Indeed, the likely drivers of seasonal differences in *W* are numerous, including convection conditions (especially driven by daytime sun angles), temperature, vegetation density differences due to seasonal leaf shedding/die back, *etc.* Just as importantly, as seasonal differences have proved to be a major determinant for human visual searcher efficacy (Chiacchia KB, Houlahan HE, 2010, as well as the current human data), use of this factor for deriving air-scent *W* values provides a single, apples-to-apples comparison with human search efforts, simplifying search managers' assessments.

The true significance of the other effectors is harder to judge. Most of them returned significant P-values. However, the smaller POD differences (≤~10% through the curves) seen with vegetation thickness and ecoregion are of indeterminate *operational,* as opposed to statistical, significance. Of course, the ecoregions sampled in this study are fairly similar subjectively; repeating the measurements in more disparate environments (*e.g.*, high-altitude mountains, warmer marshy areas, desert areas) might identify more clearly significant differences. The humidity conditions, on the other hand, correlated with no significant difference in *W* upon repeated-measures correction. This was surprising, at least among the dog-handler community, in which humidity is subjectively considered to be important for scent generation (though see the Limitations section below).

As noted in the Introduction, one study suggested that navigational errors impose errors of ~10-20% in prospective POD determinations (Perkins D, 2018). Thus, it may be that, for prospective planning at least, differences of ≥~20% may be necessary to ensure operational significance. On the other hand, acquisition of GPS tracks of teams — possible, through smartphone apps, for every searcher in a human team even if sufficient numbers of dedicated GPS units are not available — allows accurate determination of retrospective PODs despite this limitation. Note that this issue will not affect the results reported here, as the human searchers were walking a set course and the dog handlers' tracks were determined by GPS.

One could argue, with some justification, that the statistical tests used here are over-sensitive for determining operational significance. This stems, perhaps, from the nonlinear relationship between POD — which is the quantity needed to manage search efforts — and *W* — a tool for deriving that quantity. Indeed, as can be seen in these results, an intuitively large absolute difference in *W*s can produce a surprisingly small difference along the resulting POD curves. Defining operational significance, then, will likely stem from practical imperative rather than statistical significance of *W* differences.

The results in applying the POD-curve method to human visual searcher data were largely in consonance with those for the dog data, with the exception that the POD-curve-derived *W*s did not have smaller SDs than those of the crossover values for the humans. One possible explanation for this is that, while obtaining a crossover *W* for each detector suffers from the large variance in *W* inherent in the small number of DOs for individual dog/handler teams or human searchers, the latter are represented by larger Ns — *i.e.*, similarly small DOs for each individual but more individuals contributing to the mean crossover *W* value. The previous and current dog data reported here rest on 4 and 10 teams, respectively, while the smallest of the human sweep-width experiments collected data from 18 searchers, and the largest 26 searchers. More work will be necessary to determine if this is the reason for the difference, but at the very least the POD-curve method does not return *larger* SDs than the crossover method.

The curious result that the POD-curve fits to the random-search model provide larger *W* values than the (definitive) crossover method (as defined by slope>1 in the least-square plot in Figure 13), the authors believe, is tightly linked to the limitations of that model. Indeed, that model has been preferred for operational use because of its theoretical insensitivity to uneven search patterns as much as its fit to the data (Koopman BO, 1980; Charnes A, Cooper WW, 1958; Cooper DC *et al.*, 2003; Koester RJ *et al.*, 2004). In addition, though the inverse-cube model (derived from a hypothetical relationship between POD and the size of an object's image on a human retina) was primarily intended as a theoretical tool, it has been cited as nevertheless fitting well to field data in some cases (Koopman BO, 1980; Koester RJ, 2020)

The problem is that the inverse-cube model generates *W* values arguably closer to the definitive crossover values at the cost of a somewhat poorer fit to the data (reflected in its <25% score on the AICc test — smaller than that for the random-search model, though not small enough to rule it out). In particular it can overestimate POD values compared with the random-search model starting around a POD of ~60% (Figure 15). This is not far from the ~50% point at which, again, we would expect the shifting POA method to begin directing search efforts to new segments. Operationally this could be a problem when using inverse-cube, causing search efforts to be redeployed when further search in an area may be warranted.

Of course, the most likely explanation for these results is that neither model completely captures the phenomenon. The often-significant P-values for the runs tests above, for either model, further support that interpretation. Both models provide a close physical fit to the data, however, with R-square values >.8 in a majority of cases.

Smaller random-search R-square values of .264 and .555 obtained, respectively, from the fit for the brown glove at the second summer State Game Lands 203 experiment and the low-vis mannequin at the spring experiment in Mount Gretna (Table 1). This is interesting because these search objects both had extremely low *W* values (the latter is the smallest measured so far that the authors are aware of for an adult-sized mannequin). Relevant to the former, *W* values have proved difficult to obtain for low-vis clues in some cases (Chiacchia KB, Houlahan HE, 2010). Relevant to the latter, the green mannequins in that experiment had an unusually low value of *W* compared with their average maximum detection radius (AMDR, the mean distance at which a searcher is capable of seeing an item whose location the searcher is cued to, as opposed to a searcher who does not already know the location of an item, as in the *W* determinations; Koester RJ *et al.*, 2014). This apparent discrepancy between ability to *see* the search object and ability to *notice* it may be because, as reported by data loggers at that experiment, the large green rocks in that area appeared to cause the searchers to progressively ignore objects of that approximate size and color, with many close-range misses that the (cued) loggers thought should have been obvious detections. The phenomenon of attention affecting detections over and above simple visibility of an object has been documented in the literature (Koester RJ *et al.*, 2014; McClanahan S, 2021). Determining whether these observations offer any kind of window into explaining the poorer fit for these two search objects will require more work.

Taking the results reported here *in toto*, the authors would recommend the larger *W* values produced by the POD-curve fit in concert with the random-search model for operational use over the canonical crossover *W* values with the inverse-cube model (or the fit-derived *W*s using that model) because of the former's marginally greater agreement with the data. The approximately linear relationship between the random-search-fit and crossover values for *W* is reassuring on this count. Indeed, that simple relationship could allow quick conversion of previously measured crossover *W*s for use with the random-search model. Alternatively, the IDEA spreadsheet for planning sweep-width experiments contains a "Data Summary Object" sheet that, with minimal reworking, can be converted into a calculator to derive C_{50} versus POD, allowing easy re-analysis of previous sweep width experiments to provide a curve to derive *W* using any statistical package allowing nonlinear least-square fits. On the other hand, either of these approaches (larger, fit-derived *W*s with random search versus canonical or fit *W*s with inverse cube) appears to provide high-R-square agreement with the data and thus could be used operationally.

The one combination the authors would *not* recommend is using crossover *W*s with the random-search model (arguably the most common practice currently), as this may systematically underestimate POD. While intuitively this might seem attractive — a deflated POD being more "conservative" by one way of thinking because it reduces, using shifting POAs, the chance of missing a subject in areas being searched — it would also tend to cause search efforts to linger in empty areas rather than expand to new areas that might contain the search subject (Charnes A, Cooper WW, 1958). The most conservative POD values are accurate ones, the operational derivation of which the authors hope to contribute to with this report.

One objection sometimes raised to determination of objective POD via effective sweep width is that, occasionally, individual search segments may vary significantly in character from the average surrounding environment and so would not provide comparable PODs. This is likely true; but just as outlier responses to pharmaceuticals do not make their use impossible, outlier *W* values will not prohibit operational use of sweep width. Granted, it is important to understand such outliers and when and how they occur to maximize accuracy of objective POD determination. Yet this argument, if taken to its logical conclusion, given the known unreliability of obtaining POD from subjective assessments (Koester RJ, 2004), is not against use of sweep width, but any use of POD at all. Ultimately, and if the PODs reported here were not reasonably consistent — and they do appear to be — this line of reasoning calls for a completely different approach to search management, which is yet to be elaborated.

Pulling back to a more theoretical level, one would not expect the same model to capture the relationship between visual acquisition of a search object and its cognitive detection by a human searcher, *as well as* the relationships between transport of odorants in a turbulent medium, their detection by a dog's olfactory epithelium, cognitive recognition of the signal, localization of the source by the dog in that turbulent medium, and the complex psychological factors in the dog/human relationship that produce a clear signal by the dog that compels the handler to follow her to the search subject. In retrospect, the apparently detector-agnostic nature of the relationship between *W* and POD suggests to the authors that factors common to these detection methods may be rate-limiting to the search process, rather than the many differences. Future progress will be necessary to tease apart this very large question.

Limitations of the Current Study

The current study has several limitations that must be borne in mind. First, despite the authors' attempts to expand the datasets, pool data from multiple experiments, and leverage as much of the collected data as possible to derive the results above, the effort remains in the realm of "small data," and thus is vulnerable to outlier effects. Also, some of the P-values that lost their significance to repeated-measures correction may in fact represent real if weaker effects. A means for greatly expanding the data — perhaps via a smartphone application that collects data while generating a log entry for users at training events or deployments — might enable collection on a vastly larger scale. Application of the results to searchers and dog teams with different techniques or training protocols, at greater statistical robustness, and above all in areas with different climate and terrain may depend on such expanded data. The authors would be interested in working with groups or individuals capable of such software development.

Another limitation is that the relationships between the external factors measured and search efficacy are by necessity correlative. While these results help document robust correlations of weather, season, terrain, or vegetation with search results, they cannot establish causality and so must be interpreted with some caution. Uncontrolled factors and correlations between effectors that were not considered in the present analysis could bias results. In addition, the complex interrelationships between season, weather, and vegetation means that many of the factors investigated above are not independent and cannot be confidently aggregated — for example, what would *W* be for an air-scent team searching in convection condition C, in the spring, with an air temperature of 15ºC, in heavy vegetation, in ecoregion 70? The factors will almost certainly not combine in a linear fashion and so for the time being must be considered separately. One possible means past this limitation might be a Bayesian causal inference analysis (Gelman A, Meng X-L, 2004) — which, again, would require expanded data.

The limitations of measuring humidity via historical records at the nearest weather station may have played a role in obscuring any effect associated with this factor. Also, the cutoff points in the humidity categories, chosen to provide groups of similar DO counts, may not have captured transitions meaningful to air-scent detection. Along the same lines, while the freezing point of water would have been an attractive temperature cutoff for that comparison, the authors did not collect enough data below 0ºC to divide the data at that point. Also, the values derived from the POD curve for convection condition A and for ecoregion 61 rested on a very small number of DOs (20 and 35, respectively), which may not be sufficient for a reliable result (Koester RJ *et al.*, 2004).

The mannequins used for the human visual experiments are similar in visual cross-section to a supine or prone human figure and are generally considered a workable stand-in for a human subject (Koester RJ *et* *al.*, 2004). However, the possibility remains that obligate use of live humans as search subjects in the air-scent tasks resulted in an unknown bias compared with the human-searcher results. As subjects for the dog tasks wore medium- to low-visibility colors (with the handlers lending camouflage when their clothes were high-vis), the dog data may systematically under-estimate PODs for dog teams searching for subjects in high-vis colors.

Finally, while the authors have endeavored to measure effectors of search efficacy as objectively as possible, we have also confined ourselves to equipment that can reasonably be carried and managed by a searcher whose primary mission is searching and not environmental measurement. The tools for measuring convection in the field, for example, may not be precise but they are accurate — and are precise enough to determine convection category (Lavdas LG, 1976; Lavdas LG, 1997; Graham H, 1994). Such factors as season, ecoregion, and sun angle are easily obtained; others, such as humidity and temperature, were derived from historical records that, as stated above, reflected these factors measured at the nearest weather station to each search task, and thus may diverge from the true values in an unpredictable way. The vegetation determination, on the other hand, was entirely subjective and thus would profit greatly from an objective method. The authors have not yet encountered a GIS data layer that documents the character of vegetation under the canopy with good-enough granularity for this application; but such a layer, if it exists, would be invaluable. Last, the assignment of which areas to train in on a given date and which dog and handler would perform each task were arbitrary; while this was necessary to meet training goals, randomization of tasks and teams would have been more objective.

Conclusions

In this report, the authors have presented a means of deriving POD — the critical number necessary to judge search efforts and plan future efforts — objectively via the relationship between effective sweep width (*W*), a distance-quantified value specific to a given detector or group of detectors, search object, and environmental conditions, and POD as measured in the field. Making maximal use of actual detection and miss distances, this method provides the first explicit link between previously measured search efforts (via coverage, or C) and POD for prospective ground-search efforts with uncertainties at least as good as and possibly better than the definitive crossover method, and with a level of statistical sensitivity likely to be greater than that needed to measure operationally significant differences between detectors, search objects, and environmental factors. In addition, the method appears to be detector-agnostic, providing good agreement between predictions and field-measured POD for both air-scent dog teams and human visual searchers using both of the most common theoretical models. The authors recommend operational use of the slightly larger-than-canonical *W* values obtained with this method in concert with the most popular, random-search model, as providing the best fit to the data by the criteria described.

Acknowledgements

Thanks are due to Keith Conover, Allegheny Mountain Rescue Group (AMRG); Stephen McClanahan, Search and Rescue Ohio (SAR-OH) and Mountaineer Area Rescue Group (MARG); and Don Scelza, AMRG and MARG, for invaluable comments on this manuscript. Crucial discussions on SAR theory came from Robert Koester and Charles Twardy.

Thanks go as well to teammates, colleagues, and volunteers in MARG, Fayette County Sheriff's Tactical Search and Rescue, Northwest Pennsylvania K9 Search and Rescue, SAR-OH, AMRG, and elsewhere for help in setting up and hiding for air-scent dog tasks. The most profound thanks go to the dogs whose work, directly or indirectly, contributed to this study: Charlie, Cinders, Cole, Friday, Mel, Moe, Rosalyn, Rosie, Pip, Sophie, Verity, and particularly Lilly, who made us look good long before we had any possible claim to it.

We would like to thank the staff and leadership of the Pennsylvania Search and Rescue Council and the Pennsylvania SAR-EX conference for their organizational assistance and for expenses related to the study site permit at the Mount Gretna experiment. Thanks go to Jack Frost and Quincy Robe for their assistance in conducting the experiment and tearing down the course. We also thank the individual searchers at SAR-EX who participated in the experiment and their teams. This experiment was one of five supported partially or in whole by contract DTCG32-02-D-R00010 from the U.S. Coast Guard. The final development of IDEA was supported by contract HSCG32-04-DR00005 from the U.S. Coast Guard. GraphPad Software donated its Prism software.

Thanks go to the members of AMRG who participated in the new State Game Lands 203 experiment, in particular Paul Jung, whose hard work helped set up the course; AMRG, for financial support from the team's general budget in obtaining some of the disposable supplies used; as well as members of the Youth Outdoors program of Cleveland Metroparks and 4H, and the individual volunteers who came out to participate as searchers and data loggers. Particular thanks go to the Commonwealth of Pennsylvania Game Commission, for permission to conduct the experiment and for special access to the Game Lands.

About the Authors

Kenneth B. Chiacchia has been an operational SAR dog handler since he first certified with Search Dogs Northeast (SDNE) of Antrim, NH, in 1992. He is currently a dog handler and search manager with Mountaineer Area Rescue Group (MARG), Morgantown, WV, of the Appalachian Search and Rescue

Conference, as well as a firefighter/EMT with Harmony Fire District in Butler County, Pa. He is also a lead evaluator and instructor in the National Association of Search and Rescue (NASAR) GSAR program and an evaluator with the Pennsylvania SAR Council air-scent dog testing program. Ken has responded to searches in Pennsylvania, West Virginia, Ohio, and elsewhere. A professional nonfiction and fiction writer, Ken trained originally as a biochemist. He is currently the senior science writer at the Pittsburgh Supercomputing Center.

Heather E. Houlahan is a professional dog trainer, writer, and farmer who lives on a small farm near Pittsburgh. A SAR dog handler and wilderness EMT with MARG, she first certified as operational with SDNE in 1992. Heather has deployed as a dog handler in Pennsylvania, West Virginia, Ohio, New England, and elsewhere, including a 2005 deployment to Mississippi after Hurricane Katrina. She is also a lead evaluator and instructor in the NASAR GSAR program, as well as the head behaviorist for National English Shepherd Rescue of the U.S. Heather writes the “Raised by Wolves” blog, at http://cynography.blogspot.com/

They are married.

Abbreviations

AMDR: Average maximum detection radius, the mean distance at which an observer cued to the location of an object can see it.

C: Coverage, the ratio of effective sweep width times the path length of a detector within an area to its size.

C_{50}: Coverage assuming an effective sweep width of 50m.

DO: Detection opportunity, a detection or miss of a search object at a right angle (lateral) to the detector’s path.

POA: Probability of area, the estimated probability that a search object is contained within a given area.

POD: Probability of detection, the probability that a given detector will detect a given search object within an area under certain environmental conditions, assuming that object is in the area.

POS: Probability of success, the probability that a given search effort will detect a search object within a given area being searched (a “segment”). Equal to POA X POD.

POS_o: Overall probability of success, the sum of POSs over all segments being searched.

W: Effective sweep width, a distance-denominated term defining the envelope within which a given detector’s number of misses on a search object equals the number of detections outside, under specific environmental conditions.

References

Charnes A, Cooper WW. The theory of search: optimum distribution of search effort. *Management Sci.* 1958;5:44-50.

Chiacchia KB, Houlahan HE. Effectors of Visual Search Efficacy on the Allegheny Plateau. *Wilderness Environ Med.* 2010;21(4):188-201.

Chiacchia KB, Houlahan HE, Hostetter RS. Deriving Effective Sweep Width for Air-scent Dog Teams. *Wilderness Environ Med.* 2015;26(2):142-149.

Chiacchia KB. Critical Separation Versus Effective Sweep Width: Bridging the Old and New Search-and-Rescue Worlds. *Wilderness Environ Med.* 2020;31(1)44-49.

Cooper DC, Frost JR, Robe RQ. *Compatibility of Land SAR Procedures with Search Theory.* Alexandria, VA: Potomac Management Group Inc; 2003.

Gelman A, Meng X-L (eds). *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*. West Sussex, England: John Wiley & Sons Ltd; 2004.

Graham H. Probability of detection for search dogs or how long is your shadow? *Response Magazine*. Winter 1994;1-7.

Koester RJ, Cooper DC, Frost JR, Robe RQ. *Sweep Width Estimation for Ground Search and Rescue.* Alexandria, VA: Potomac Management Group Inc; 2004.

Koester RJ. *Lost Person Behavior.* Charlottesville, VA: dbS Productions; 2008.

Koester RJ, Chiacchia KB, Twardy CR, Cooper DC, Frost JR, Robe RQ. Use of the visual range of detection to estimate effective sweep width for land search and rescue based on 10 detection experiments in North America. *Wilderness Environ Med.* 2014;25(2),132-142.

Koester RJ. Enhancements to Statistical Probability of Area Models based upon updated ISRID data collection for Autistic Spectrum Disorders and Typically Developing Children. *J Search Rescue.* April 2020;4(1):35-63.

Koopman BO. *Search and Screening: General Principles with Historical Applications.* Elmsford, NY: Pergamon Press; 1980.

Lavdas LG. A groundhog's approach to estimating insolation. *J Air Pollution Control Assn.* 1976;26:794.

Lavdas LG. *Research NoteSRS-4: Estimating Stability Class in the Field.* Juliette, GA: U.S. Department of Agriculture; 1997.

McClanahan, S. Object recognition and detection: Potential implications from vision science for wilderness searching. *J Search Rescue.* July 2021;5(1):1-17.

Perkins D. *The Variable Nature of Probability of Detection for Ground Search Teams.* Ashington, Northumberland, UK: The Centre for Search Research; April 2018.

Syrotuck WG, Syrotuck JA. *Analysis of Lost-Person Behavior.* Mechanicsburg, PA: Barkleigh Productions; 2000.

Optimizing Wilderness Search and Rescue: Discovery and Outcome

D. Kim Rossmo PhD
School of Criminal Justice and Criminology, Texas State University

Lorie Velarde MSc
Irvine Police Department

Thomas Mahood MSc
Formerly with Riverside Mountain Rescue Unit

USA

Email krossmo@txstate.edu

Abstract

This article is a follow-up to a 2019 analysis that applied Bayesian probability techniques to the search for a missing hiker in Joshua Tree National Park. In February 2022, that hiker was found. Here, we compare the location of his remains with the results of our prediction model and discuss further implications for optimizing wilderness search and rescue.

KEY WORDS: *Wilderness Search and Rescue, Lost Persons, Bayesian Analysis, Resource Optimization*

Introduction

In 2019, the *Journal of Search and Rescue* published an article in which we discussed the use of Bayesian methods to support wilderness search and rescue (WiSAR) efforts (Rossmo, Velarde, & Mahood, 2019). The case of William Ewasko, a hiker who went missing in Joshua Tree National Park (JTNP), was used to illustrate how different types of spatial evidence could be integrated into an optimal probability search map.

Ewasko disappeared in late June 2010. Twelve years later, in early February 2022, hikers discovered human remains and a backpack, with Ewasko's identification, in the northwest section of JTNP (Hazel, 2022). The San Bernardino County Sheriff's Department later confirmed his identity through dental records. In this follow-up article, we discuss the accuracy of our original prediction and further implications for WiSAR Bayesian analysis.

William Ewasko's Disappearance

William Ewasko was an experienced hiker who regularly visited Joshua Tree National Park in Southern California (see Figure 1). On June 24, 2010, he set off on an ambitious day hike, planning to be finished around 5:00 pm (Mahood, 2012). While arguably not wild or remote, the 1,240 square mile (3,213 km^2) park can still be dangerous, with large rocks, deep canyons, and hot summer temperatures (Manaugh, 2018).

Figure 1: Joshua Tree National Park

When Ewasko failed to call in that evening, park rangers checked the trailheads on his itinerary. On June 26th, a California Highway Patrol helicopter spotted his vehicle at the Juniper Flats Trailhead parking area. The search then focused on locations reachable from here, including Quail Mountain. At 6:50 am, Sunday, June 27th, Ewasko's mobile phone registered ("pinged") with a cell phone tower on Serin Drive, Yucca Valley, just to the northwest of JTNP, indicating he had moved well beyond Quail Mountain. Searchers redeployed in response but were unable to find any trace of the missing man.

The official search was called off on July 5th, though numerous follow-up searches were conducted over the years by experienced volunteers. The electronic GPS data in the JTNP Ewasko master file eventually totaled 1,772 person-miles (2,852 km) of official and volunteer search tracks (see Figure 2).

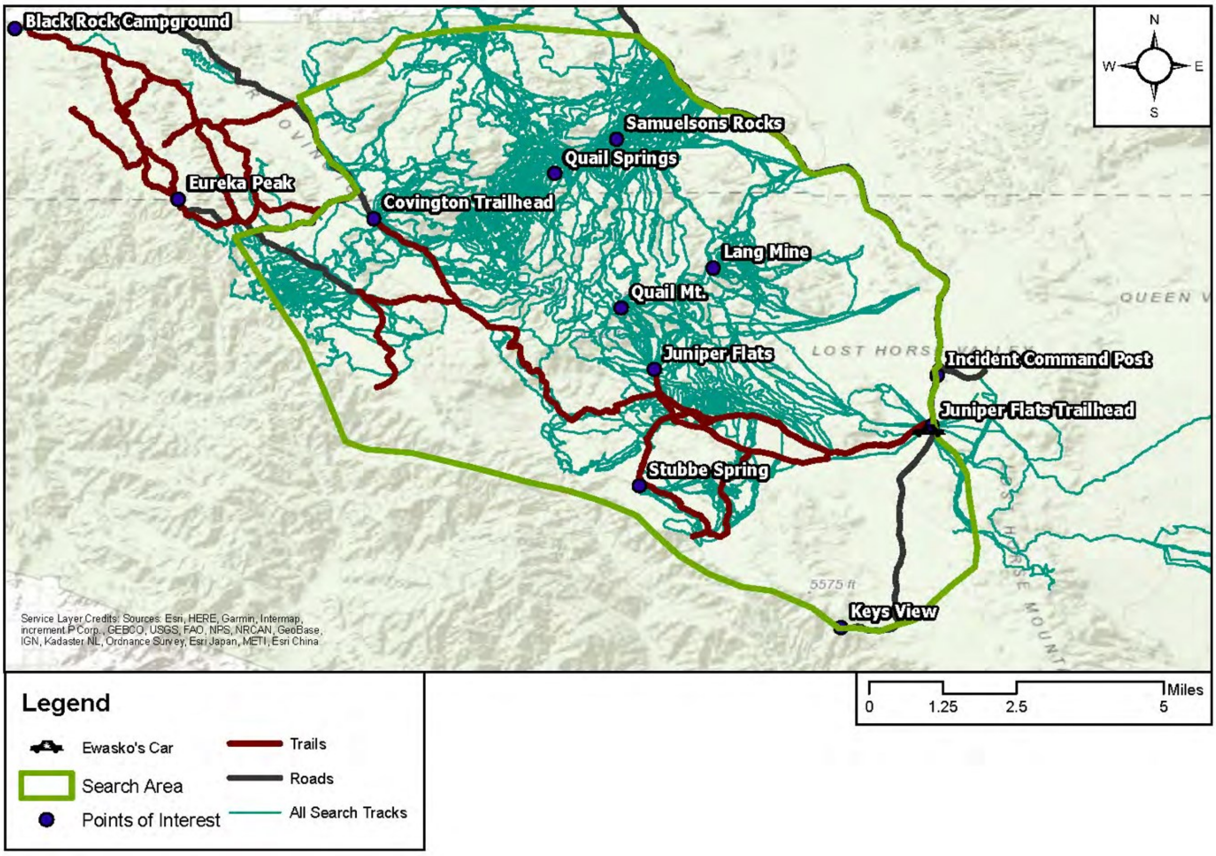


Figure 2: Search Tracks, Trails, and Roads

Bayesian Optimal Probability Analysis

The purpose of our original analysis was to demonstrate how Bayes' theorem can be used to generate priority search maps for WiSAR by combining different evidence sources – in the Ewasko case, the location of a cell phone tower, the range of its ping (approximately 10.6 miles, or 17.1 km), a terrain altitude viewshed analysis, and prior search tracks (Eddy, 2004; Iversen, 1984). Further details, including a full description of the Bayesian analysis, can be found in Rossmo, Velarde, and Mahood (2019).

Figure 3 shows our resulting optimal search map. Colors, ranging from blue to red, represent the various probabilities for finding Ewasko (the probability color scale in the legend lists relative numbers, not specific probabilities). The non-coterminous peak areas are all situated to the north of Quail Mountain. The inset map depicts the zone of highest probability.

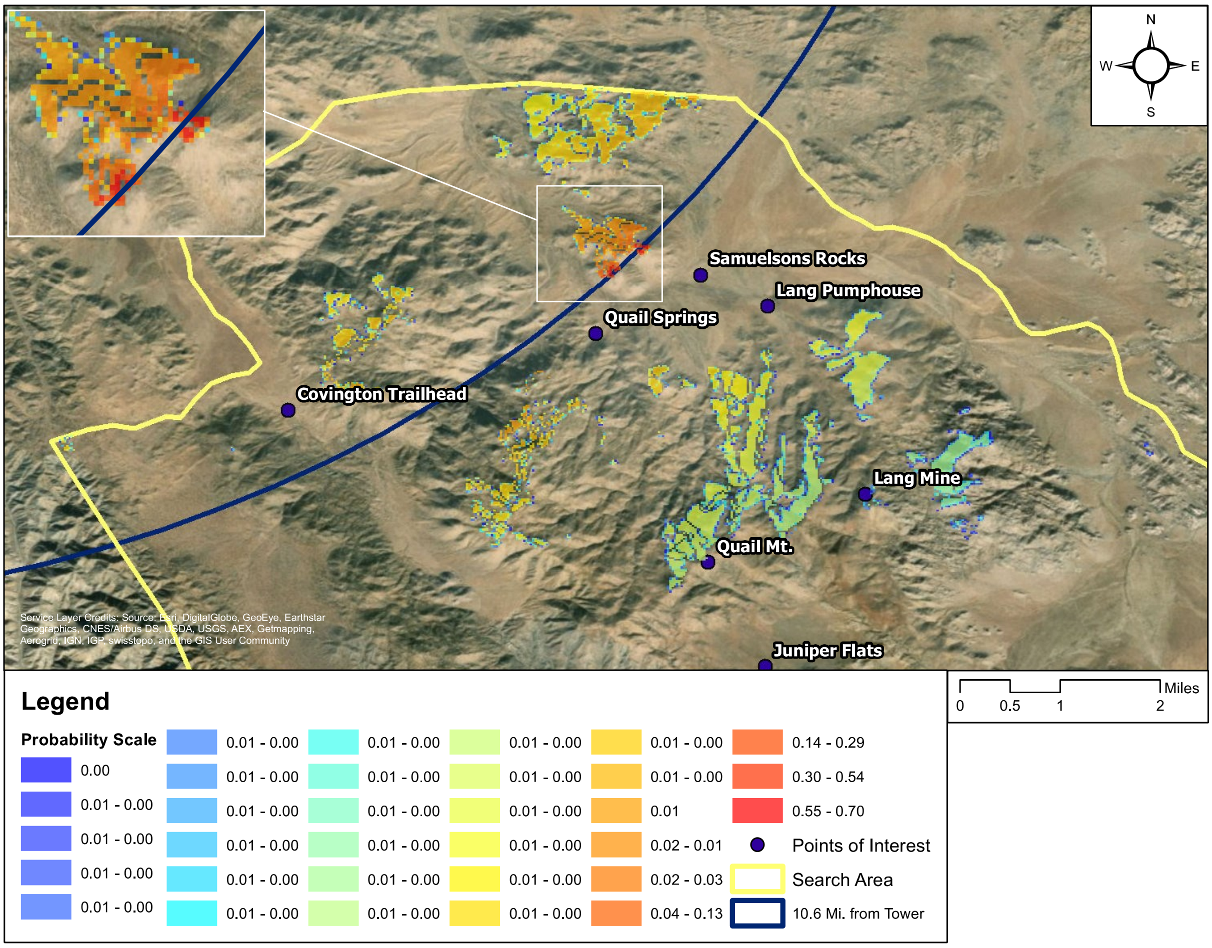


Figure 3: Optimal Probability Search Map

Optimal search areas, ranked by color scale, shown in relationship to JTNP landmarks and 10.6-mile ping radius from Serin Drive cell phone tower. Inset box at upper left magnifies the highest probability region.

Location of Remains

Ewasko's remains were discovered on a saddle along a ridge spine (see Figures 4 and 5). This spot was about 3.9 miles (6.3 km) north of Quail Mountain, his presumed destination, and 7.6 miles (12.2 km) northwest of his parked vehicle. Sadly, he was only about 1½ miles (2.4 km) from Park Boulevard when he succumbed, close enough that vehicles could be heard, though not seen.

Figure 4: Area of Ewasko Remains

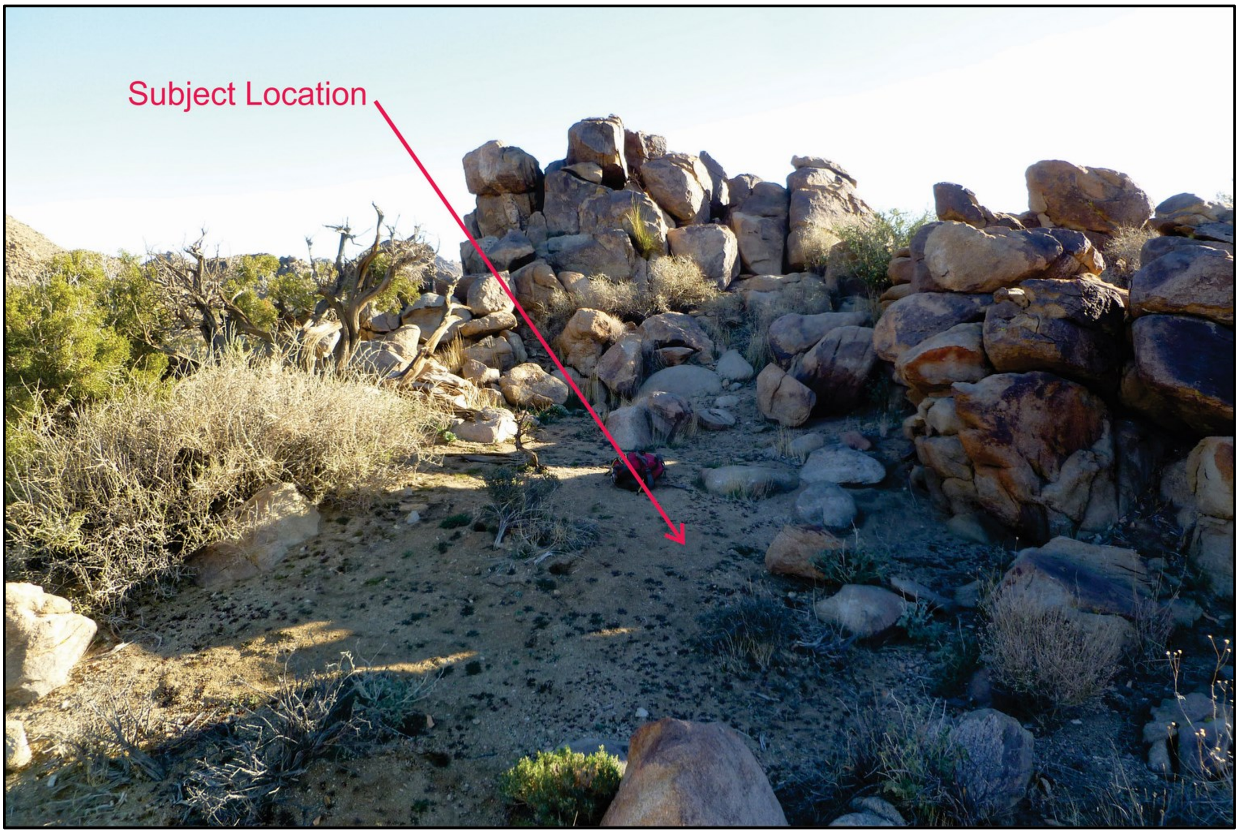


Figure 5: Location of Ewasko Remains

Figure 6 shows the locations of Ewasko's remains and his parked vehicle, overlaid on a map of the peak optimal search areas from our original analysis. The region of highest probability fell along the 10.6-mile ping radius from the Serin Drive cell phone tower, north-northeast of Quail Springs (see inset box Figure 3). Ewasko's remains were found just over a half-mile (0.83 km) northeast of this area, and only 0.28 miles (0.45 km) southeast of the nearest identified high probability area.
The distance from this spot to the Serin Drive cell phone tower is 10.25 miles (16.5 km). The exact location lacks cell phone coverage, but if Ewasko approached from the south, consistent with his travel direction, he would have passed through a small area of coverage close to the 10.6-mile ping radius. Given the temperature and three-day time lag, it was not expected he would have been able to walk much farther.

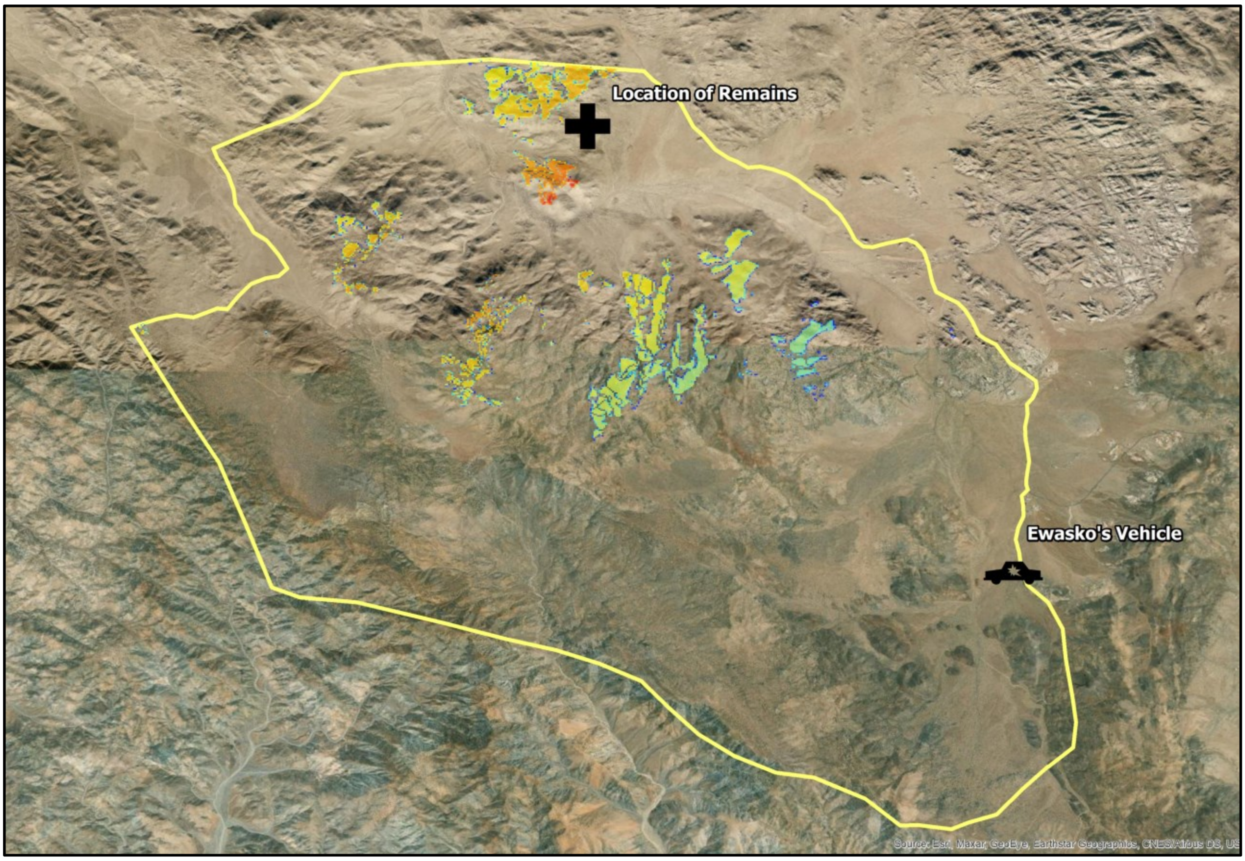


Figure 6: Optimal Probability Search Map and Locations of Ewasko Remains and Vehicle

---

Discussion

The outcome of the Ewasko case illustrates the importance of remembering that any analysis in a search and rescue effort is time dependent. While there was no cell phone coverage where Ewasko's remains were found, the viewshed analysis, park terrain, and tower ping radius all indicate he was headed north and passed through the peak probability area displayed inside the Figure 3 inset box. He was able to walk another half-mile before expiring. The possibility that a missing person will still be mobile must always be considered. Terrain and intent remain important considerations.

All of the assumptions in our original analysis appear to have been valid. There are only three possible paths between his parked vehicle and where he was found. While terrain constraints, cell phone coverage limitations, and timing suggest a few possible theories, unless Ewasko took notes or photographs we may never know for certain his exact hiking route or decision rationale. At best guess, it appears he was the victim of a lack of local knowledge regarding the difficulty of the specific area of JTNP he was hiking through, curving and disorienting canyon bottom trails, limited visibility, hot summer temperatures, and eventual dehydration. Circumstances suggest he may also have suffered an ambulatory injury, such as a bad sprain.

Recently, two of the authors were asked by Texas Search and Rescue (TEXSAR) and local law enforcement to assist in the search for a missing person. Optimization techniques similar to those used in the Ewasko case were applied to the effort. While there were no phone pings in the Texas case, the clothing of the missing person was discovered, while the road, fencing, and landscape provided some constraining influences on his movement. Data from the International Search & Rescue Incident Database (Koester, 2008) helped calibrate the distance-decay functions used in that analysis.

Tens of thousands of photographs of the relevant area were taken using drones (UAS, or unmanned aerial systems). Equipped with sensors for different spectral wavelengths, UAS can help locate human remains (Wescott, 2020). When bodies begin to decompose, the activities of bacteria and insects can generate enough heat or thermal energy to make the bodies detectable by infrared sensors. If the body has been scavenged and scattered, the near-infrared spectrum can detect the organically rich soil produced when bodily fluids leak during decomposition. Ultraviolet light is ideal for detecting bones as they fluoresce. If drones have been employed in a WiSAR effort, the Bayesian optimal probability search map provides an evidence-based strategy for prioritizing the review and inspection of large numbers of UAS photographs.

The Texas investigation remains unresolved, and may remain a mystery because of the high level of local carnivore animal activity (Haglund, 1997). This region has coyotes and feral hogs, the latter a problematic invasive species capable of eating large bones. However, the case does illustrate the flexibility of the Bayesian approach and its utility in managing information overload challenges.

Conclusion

The failure to find Ewasko was surprising as JTNP is lightly forested and its terrain is not difficult to survey (see Mahood, 2016, 2018). At the time of publication of our original article, there were no remaining viable theories as to his whereabouts. Subjects who are not quickly found by search and rescue personnel have usually done something unanticipated, such as traveling further than estimated, leaving the designated search area, or heading in an unexpected direction. All three of these happened in the Ewasko case.

Mahood, one of the coauthors of this article, has extensive search and rescue experience in Southern California. He initially believed it unlikely Ewasko would be found in the peak probability areas at the north of the Bayesian map as there was no known reason why he would climb there from Quail Wash. Except it turned out there was. And, while we don't know why, Ewasko must have considered it important.

It has been consistently shown that algorithms outperform human judgment (Kahneman, Sibony, & Sunstein, 2021). Despite the accuracy of the prediction here, however, more research and further testing is needed, and we recommend considering the two approaches as complimentary. What is indisputable is that as a search progresses, and as more and more areas become cleared, the unlikely becomes more probable. It is worth repeating a quote from Pete Carlson, Riverside Mountain Rescue Unit: "If you haven't found them, then they're someplace you haven't looked yet" (Manaugh, 2018).

About the authors

D. Kim Rossmo is the Director of the Center for Geospatial Intelligence and Investigation at Texas State University. He has researched and published in the areas of geographic profiling, environmental criminology, and criminal investigative failures. He received his PhD in Criminology from Simon Fraser University.

Lorie Velarde is a geographic information systems analyst with the Irvine Police Department in California. She has a Master of Science in Criminology, a Bachelor of Arts in Social Ecology, and a California Department of Justice Certification in Crime and Intelligence Analysis.

Thomas Mahood is a retired civil and traffic engineer, formerly with the Riverside Mountain Rescue Unit in California. He holds a Bachelor of Science in Engineering from California State University, Long Beach, and a Master of Science in Physics from California State University, Fullerton.

Abbreviations

GIS	geographic information system
GPS	global positioning system
JTNP	Joshua Tree National Park
WiSAR	wilderness search and rescue

References

Eddy, S. R. (2004). What is Bayesian statistics? *Nature Biotechnology*, *22*, 1177-1178.

Haglund, W. D. (1997). Dogs and coyotes: Postmortem involvement with human remains. In W. D. Haglund & M. H. Sorg (Eds.), *Forensic taphonomy: The postmortem fate of human remains* (pp. 367-381). Boca Raton, FL: CRC Press.

Hazel, D. (2022, March 2). Enjoying nature: A 12-year mystery is finally solved. *Crossville Chronicle*. Retrieved April 7, 2022, from https://www.crossville-chronicle.com/news/glade_sun/enjoying-nature-a-12-year-mystery-is-finally-solved/article_18f61480-9a47-11ec-b795-f7115d76f958.html.

Iversen, G. R. (1984). *Bayesian statistical inference*. Sage university paper series on quantitative applications in the social sciences, 43. Beverly Hills: Sage.

Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A flaw in human judgment*. New York: Little, Brown Spark.

Koester, R. J. (2008). *Lost person behavior: A search and rescue guide on where to look – for land, air and water*. Charlottesville, VA: dbS Publications.

Mahood, T. (2012, February 22). *Searching for Bill Ewasko* [Blog post]. Retrieved from http://www.otherhand.org/home-page/search-and-rescue/searching-for-bill-ewasko/.

Mahood, T. (2016, May 14). *Ewasko thoughts, May 2016* [Blog post]. Retrieved from http://www.otherhand.org/home-page/search-and-rescue/searching-for-bill-ewasko/ewasko-thoughts-may-2016/.

Mahood, T. (2018, August 4). *A Bayesian approach to looking for Bill Ewasko, August 2018* [Blog post]. Retrieved from http://www.otherhand.org/home-page/search-and-rescue/searching-for-bill-ewasko/a-bayesian-approach-to-looking-for-bill-ewasko-august-2018/.

Manaugh, G. (2018, March 22). Tragically lost in Joshua Tree's wild interior. *The New York Times Magazine*. Retrieved July 13, 2019, from https://wholeworldlatestnews.wordpress.com/2018/03/26/tragically-lost-in-joshua-trees-wild-interior-by-geoff-manaugh-2/.

Rossmo, D. K., Velarde, L., & Mahood, T. (2019). Optimizing wilderness search and rescue: A Bayesian GIS analysis. *Journal of Search and Rescue*, *3*, 44-58.

Wescott, D. (2020). Using drones to locate and document outdoor crime scenes. *Faculty Research Spotlight*. San Marcos, TX: Office of Research and Sponsored Programs, Texas State University.

Postmortem submersion intervals in the River Thames

Ian Greatbatch PhD[a], David Livingstone[b], Gordon Hunter PhD[b] and Yvonne Saunderson[c]

[a]School of Earth & Environmental Sciences, University of Portsmouth, UK; [b] Faculty of Science, Engineering and Computing, Kingston University London, UK, [c] National Health Service, UK

Email: Ian.greatbatch@port.ac.uk

Abstract

This paper catalogues a dataset drawn from Metropolitan Police records in the period 2004-2015 regarding drowning victims recovered from a tidal stretch of the River Thames and provides a comparative study with a similar dataset from the time period 1956-1959 by the County of London (Western District), H.M. Coroner of that time, Gavin Thurston, from the information gathered in his professional role. In addition to drawing comparisons between these time periods the paper draws inferences regarding the Post-Mortem Submersion Interval (PMSI) and proposing further study required. Both datasets show a significant bias towards male subjects counter to other comparable data in the literature. This bias is even more pronounced in the 2004-2015 period where additionally there are significantly more incidents during Full Moon Lunar Phase. Some weak seasonal trends were observed regarding equinoctial peaks but these were not statistically significant.

There are few clear trends observable in factors that might influence Post-Mortem Submersion Interval although there was a weak and counter-intuitive inverse relationship between clothing weight and time in water.

KEY WORDS: PMSI, Drowning, GIS

Introduction

This paper concerns bodies recovered from the tidal reaches of the River Thames during the period 2004-2015 and in particular their Post-Mortem Submersion Interval (PMSI) (the interval between date of entry and date of recovery). The cases included are those where the known cause of death was by drowning and both the entry and exit points of the bodies were known by the Metropolitan Police's Marine Policing Unit (MPU). We define PMSI in a literal sense, meaning the time spent submerged, and also analyse the distance travelled, but do not consider rates or levels of decomposition.

The key work concerning body recovery from the Thames in London is Thurston (1960), which considered the cases of bodies recovered between 1956-1959 and there has been no work since 1960 concerning the movement of bodies in the Thames. This work uses technology and methods unavailable to previous authors, and addresses a considerable gap in understanding concerning preventable death in a global megacity and European Capital.

The general structure of this paper follows to some extent the format of Thurston 1960 and, where relevant, comparisons are made with these findings. The rationale for this was that the Thurston paper represents the *only* work concerning this aspect of suicide in this major city, and it may be of value to compare results between the two periods, and perhaps determine what factors, if any affect rates and physical geographical characteristics of suicidal drownings in the Thames.

The Thurston paper presents information regarding 82 cases in the period 1956-59, over twice the number of cases in the dataset used in this work. Thurston offered a set of descriptive statistics, illustrated by some case studies, and, where relevant a comparison of the 2004-2015 data is made, and differences are analysed. For the Thurston dataset only secondary data was available and this was not universally applicable to all 82 cases. Therefore, in some circumstances a direct comparison cannot be made, this is discussed in the relevant sections of this paper. The following sections will discuss elements of the river that may have changed (tidal flow, turbidity, development, SAR provision etc) before moving on to the main data analysis of the paper.

Since Thurston's work there has been no published investigation of the patterns and properties of drowned bodies recovered in the Thames in London. The river, its immediate environs and London in general have all changed considerably, as have the organisational structure of the emergency services response on the river. This work will use a dataset compiled by the MPU and a Geographical Information System (GIS) approach to quantify and disseminate the available data on drowning in the tidal Thames. Additionally, this work will seek to provide a baseline set of data concerning the Post-Mortem Submersion Interval (PMSI) of bodies in the Thames, as well as observations on the movement and descriptive statistics of the victims.

The Search and Rescue (SAR) provision on this stretch of the Thames is one of the most drastic changes over the period of this study. The SAR provision was initiated by the 2000 Lord Justice Clarke's Thames Safety Inquiry, in response to the 1989 *Marchioness* disaster on the Thames, in which 51 people lost their lives (Butcher, 2010). The impact of a more efficient rescue system on the river is likely to have had a significant effect upon the death rate of suicide attempts by increasing the number of unsuccessful attempt, and reducing the number of deaths. There had previously been a Marine Accident Investigation Branch report (1991), the Hayes Report (1992) and a Coroner's Inquest (running from 1990-1996). The Inquiry made a number of additional recommendations, amongst which were a consultation on the consumption of alcohol by people in charge of vessels, funding for a formal safety assessment of search-and-rescue facilities on the Thames; and funding for experimental life-saving equipment at locations along the Thames (Butcher, 2010).

As of 2003, Her Majesty's Coastguard (HMCG) began coordination of all Search and Rescue (SAR) activity on the river. Four new Royal National Lifeboat Institution (RNLI) lifeboat stations were established, with three of those four being permanently crewed 24 hours a day, and one relying on a pager system to mobilise locally-based crew members. In addition, the Port of London Authority (PLA) also has Harbour Master and other inspection vessels,and the London Fire Brigade has one permanently afloat fire boat. All of these resources have some SAR capability and can be called upon for search and rescue operations and body recovery. However, the main responsibility for body recovery lies with the MPU and they would be the most likely resource to be mobilised for this operation on the tidal Thames.

Background

River Thames Environment

Below Teddington, the river is tidal, with a large tidal range and strong flows, with the potential to move large objects some distance but is locked and managed above that. There is a lock and weir at Richmond (within the tidal portion of the Thames), which also serves to control the flow and fluvial output into the tidal river. The Thames flood barrier at Greenwich protects London from North Sea storm surges, but also clearly has an impact on the flow of the river. Considering changes since Thurston's analysis, Fig 1 shows a small decline in flow rate, but changes to the nature of the river (a move away from predominantly industrial use) are likely to have had an impact on turbidity and composition of suspended loads further downstream (Werner, 2015).

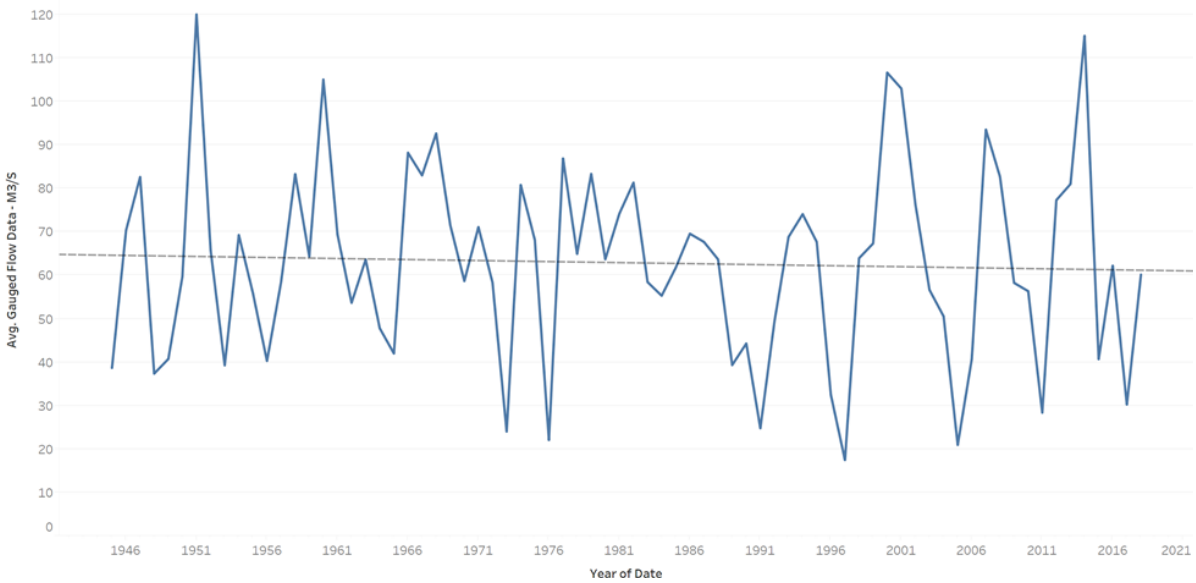


Figure 1: Average Daily Flow (cubic metre per second) from gauged data at Kingston Upon Thames (1945-2018)

The amount of water in the river is related to the rainfall inundating the river's catchment and as such it can be seen that there is a period from October to January where there is typically higher rainfall, more water in the river system and resultant faster fluvial flows (Fig 2). This will be likely to have an impact on the tidal action, on any suspended load, but it is difficult to be specific about the nature or extent of that impact (UK Met. Office, 2020).

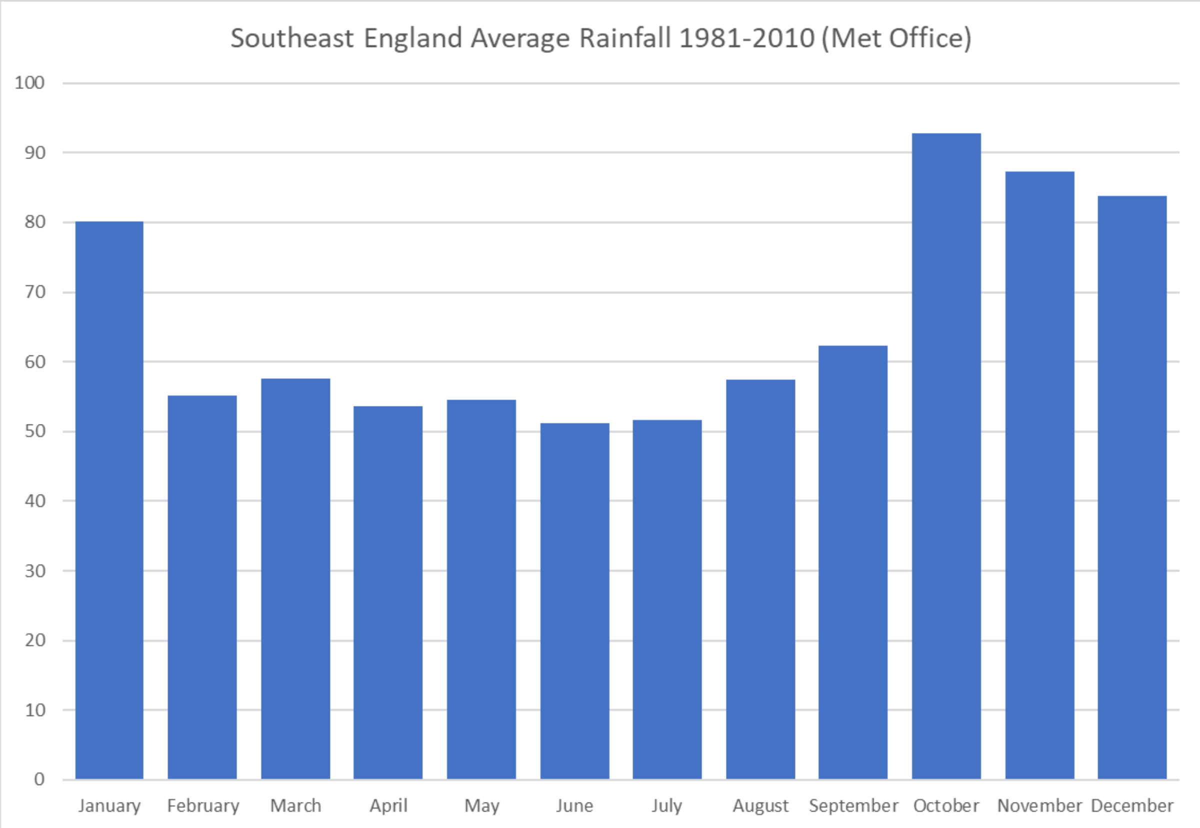


Figure 2: Average Monthly Rainfall (millimetres) for South East England 1981-2010

River Thames: Suicides and Drowning

Drowning itself is a major cause of death, disability, and loss of quality of life and is a leading cause of death among children globally (Branche and van Beeck, 2014). According to Martyn, (2014), there were nearly 400,000 deaths globally in 2004, with over 90% of them occurring in low to middle income countries. The UK, has a relatively low rate of fatal drowning, well under 1 death per 100,000 population (in comparison to Moldava, with 7/100,000) (Saxena *et al.*, 2014) and the highest rate of drowning in swimming pools in the UK is 14-15 year olds during summer months (ROSPA, 2007). Across Europe, suicide by drowning represents less than 5% of all deaths by suicide, with this rate being higher among females (with some countries showing female rates of over 10% suicides by drowning (Varnik *et al.*, 2008)). This may be due to ease of access to water, which of course varies geographically, or as a function of the nature of the suicide attempt. For example, some individuals may select an isolated location, to reduce the likelihood of rescue (and in some cases fill pockets with stones, or bind their hands), some individuals may select a well-populated location, to ensure rescue, where the intention is to attract attention, but not actually die (Lunetta and Connolly, 2014).

Concerning suicidal drowning in general, Copeland (1987) provided an analysis of both location, demography and rationale, determining that canals or the home were the most frequent locations (both 21.4%). The study determined that the population predominantly consisted of older white males, and was typically as a result of some depression concerning a variety of factors (Copeland, 1987).

Byard et al (2001) carried out a retrospective analysis of South Australian suicides by drowning, looking at 176 cases that dated from 1980 to 2000, although only 29 of these related to those who drowned on a waterway (as opposed to a bath, swimming pool etc). There were 76 males and 47 females, with a total age range of 16 to 88 years. Female victims tended to be older than male victims, and victims were more likely to be suffering from mental illness than from a physical illness.

There are fifteen bridges over the Thames in central London, and whilst being of great utility to the city, and in some cases serving as tourist attractions in their own right, they have sadly and inevitably been used by those seeking to take their own, or others' lives. There are parapets and railings along all of the bridges, but the requirement for them to afford views of the river and the city mean that they are not impassable. In the last 50 years the river has been cleared of industrial buildings considerably along its banks with both north and south banks offering access to the river along almost all of their length (Haywood, 1998; Ward and Robinson, 2000). The HMCG control centre is based at Greenwich, but has live links to CCTV cameras on the majority of bridges over the Thames, as well as other sections of the river bank. London also has a plethora of "passive" cameras around the Thames, which could be accessed for forensic investigation after an event (Skogan, 2019).

The fluvial flow of the Thames does have some seasonal variations (above), and there is some evidence for seasonality of affective disorders, with Lukmanji et al. (2019) finding a significant seasonal variation, amongst certain age groups, of feelings of depression or low self-worth. However, this same study found that despite this, there was no increase in suicidality in any season. Rotton and Kelly, (1985) and Greatbatch et al. (2019) also found no statistical increase in incidents for the emergency services, psychiatric admissions, homicide, disasters or suicide during full moon periods.

Newiss and Greatbatch, (2019) found a significant number of men missing from a night out were found dead in water, although none of these were suspected of being suicides. Partonen et al. (2004) found that suicides in Finland were not evenly distributed across seasons, were higher around the phases of a new or full moon, and attributed this mismatch to changes in ambient temperature and light. Lasota et al. (2019) found a spring peak of suicides was for men, and an autumn peak for women, and Aguglia et al. (2019) found that high-lethality suicide attempts were significantly greater than low-lethality attempts during the months with most sunlight (June and July). In conclusion, the literature does not provide a conclusive answer as to the likelihood of a seasonal affect in suicide by drowning in the Thames.

Post-Mortem Submersion Interval

Detailed research into the precise mechanics and movement in water of the human body is however difficult to carry out. This is due in part to the unpredictable nature of drowning incidents, along with the likelihood of recovery by the authorities as soon as a body is discovered. It is also impractical and unethical to carry out experiments with trackers on either dummies or cadavers, for fear of upsetting the local population. In this sense, water based forensic experimentation is at a disadvantage to terrestrial research, where body farms can be used to test factors and circumstances in distribution of bodies by scavengers and general decomposition (Williams *et al.*, 2019).

There is limited research on the level of clothing remaining on drowned victims, but Helmus et al. (2018) found with experiments on cadavers in private rivers that it was common for bodies to remain at least partially clothed after immersion, although the focus of their study was focused on the persistence of DNA on clothing. Dennison-Wilkins (2021) found that questionnaire returns from recovery teams suggested that differences in buoyancy relating to clothing and footwear existed. The differences were related to the amount of clothing remaining or initially present on the body and the type and construction of clothing. They also found that buoyancy appears to increase with age of the individual and that the manner of death itself has an impact on buoyancy, which logically then impacts upon PMSI.

According to Haw and Hawton (2016), most bodies drowned as a result of suicide were found wearing outdoor clothing, with a smaller group in nightwear, swimwear or naked. Lunetta et al. (2014) note that post-mortem identification of bodies is often difficult as “clothing and other personal effects are lost in water”. Armstrong and Erskine, (2011) note that nude recovered bodies may be as the result of the “stripping off of clothing caused by rough water conditions or snagging against underwater debris versus a dumped body with death by other means”. It therefore may be that clothing is stripped off a body based on a function of flow conditions, river bed roughness, clothing material and length of time submerged. The stretch of the Thames considered in this study is not uniform in terms of flow, tide strength or river bed rugosity. However, it is certainly true that all of the conditions that may contribute to clothing being stripped post-mortem do exist to some extent in all reaches.

The length of time that bodies spend underwater, once submerged, is of importance to later investigations, and determining the events leading up to death, especially so in any criminal investigation. As a result, there have been a number of attempts to generate heuristics that relate the state of decay of a body to submersion interval, and other environmental factors, either using animals as human analogues (Humphreys *et al.*, 2013), or through data analysis from fatality databases. van Daalen et al. (2017) used the date a victim went missing as a start or immersion date, and used that (with recovery date) to estimate PMSI and investigate levels of association between PMSI and decomposition. The mean PMSI was 41 days, based on 38 cases of bodies recovered from the North Sea by the Dutch Department of Missing Persons in the North Sea (BVPN).

The movement of drowned bodies through water, and the various related characteristics of that movement are also of interest to forensic investigators, leading potentially to the identification of missing people or leading searchers to locate drowned victims faster and more efficiently (Byard, 2018; Mateus *et al.*, 2013; Reijnen *et al.*, 2018)

Clearly, functions of water movement, through currents or flows and a body's specific gravity (SG) (combining physical characteristics, levels of gas within the body and additional items such as clothing) determine how fast, in what direction and at what depth a body is suspended. With a specific gravity less than 1.000, a body will sink in fresh water and remain there until gases within the body are produced as a result of decomposition, bloating occurs and the SG is raised (Lucas *et al.*, 2002). Water temperature, body fat, river microbial activity, weight of clothing and the roughness of the river bed are all factors in any post-mortem movement, as they affect the speed and level of bloating, as well as the forces acting upon the body resulting in movement, and as previously started, all of the environmental conditions impacting body movement are present within the tidal and non-tidal reaches of the Thames in London.

Recent research into PMSI in the UK has been focused within the forensic science field, and the prediction of PMSI from post-mortem injuries on bodies recovered. Heaton et al. (2010) found that a combination of submersion and temperature had a measurable and predictable impact on the level of decay to the body, and as a result made some progress towards being able to predict the duration of submersion from the level of decay (ibid). They reviewed 18 cases from the Rivers Clyde and Mersey in the UK, and had access to closed case files, autopsy results and crime scene reports to determine the level of decomposition.

Method

The data regarding the drowned victims was provided under license from the National Crime Agency, the National Missing Persons Bureau and from the Metropolitan Police, although the data was specifically sourced by the MPU. The criteria for records to be forwarded were that both the recovery location *and* point of entry were known. There are records kept by both the RNLI locally, and the police, of all bodies recovered, which on average numbers between 20 and 30 a year. Records where both entry and recovery are known are, understandably, fewer in number, as it is rarer for someone to be observed entering the water and not be rescued immediately. Some of these records overlap with the dataset used in this study, but there is no requirement for harmonised recovery records. The contemporary dataset will be referred to at the Met. Police 2004-15 dataset in this work.

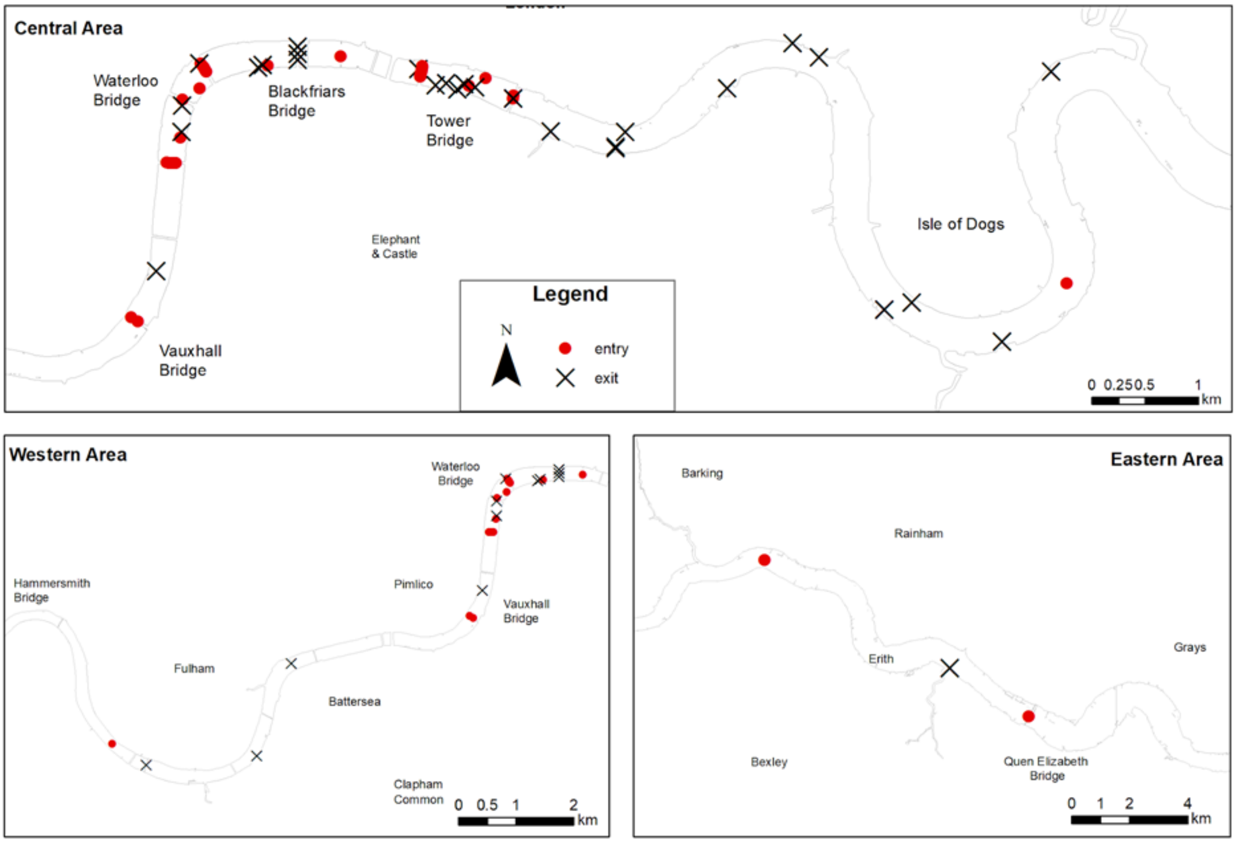


Figure 3: Entry and Exit Locations of Drowned Bodies (June 2004-May 2015) from Metropolitan Police Records

The data contained 33 records of drownings in total for the period covering June 2004 to May 2015, the distribution of which are shown in Figure 3. The dataset contained the sex, age and clothing of the casualty when recovered, as well as the date and location of both entry and recovery. All cases were all believed to be suicides (apart from 1 case, which was accidental), and all but one casualty was male.

Standard data processing allowed additional data to be incorporated to the dataset, such as phases of the moon, and PMSI (in days). In addition, the level of clothing remaining on each recovered victim was quantified using a simple grading system as shown in TABLE 1. The number of items of clothing was added as an integer to the database to allow descriptive statistics and comparisons to be produced.

Clothing Level	Score
Naked, or near naked	1
Clothing but no shoes	2
Clothing and shoes or trainers	3
Clothing and heavy boots, or heavy overcoat	4

Table 1: Coding for clothing levels.

The entry and recovery points were digitised in ESRI ArcMap 10.2.2 to as close a position as could be determined from the description (Fig. 3), although some uncertainty will be present in the data as a result of there being no record of where exactly from a bridge or jetty a person jumped, nor where precisely they were recovered. The description of recovery tends to be a guideline location on the river, denoted by a local landmark (such as Chelsea Harbour Pier) which allows for a point to be placed in space in order to carry out analysis, but is inherently uncertain.

The points were then joined by lines that ran along the centre line of the river (having no other information as to their true track) and these were snapped to the centre line (rather than using a straight line that would have cut across the *meanders* of the Thames) using a form of *Dynamic Segmentation* (Chiou *et al.*, 2010) and the distance along the river from entry to recovery points was calculated.

Results

The key text concerning drowned body recovery in London is Thurston's 1960 paper, which considered cases of bodies recovered from the Thames between 1956-1959. There were 66 body recovery cases in that period. In 12 cases where the victim was seen to enter the water, however in 21 additional cases corroborative evidence for entry point was available such as recovered personal belongings or discarded clothing. This resulted in 33 of the 66 cases available to Thurston, having known entry and exit points. Nearly all cases were considered suicides even though 65% of them had open verdicts. The proportion of drowning cases in general, and suicide cases in particular available to Thurston as Chief Coroner of the County of London (Western District) are likely to be greater than the proportion of those available to the Metropolitan Police for the reporting period covered in this paper and so it is not possible to infer anything about the rate of suicide by drowning by comparison of the numbers recorded during the respective periods.

What follows are a set of comparative descriptive statistics between the Metropolitan Police cases (2004-15) and Thurston's cases (1956-59) following the format adopted by Thurston, and illustrated by some case studies. The geographical area for the 1956-59 Thurston cases coincides with those entry and exit points falling entirely within the Western Area of the Metropolitan Police cases 2004-2015 (Fig 3), and these number a total of 12 out of the 33 cases. The statistics are further subdivided according to the Western Area to allow direct comparison with the Thurston cases. Where relevant the Metropolitan Police dataset is split into geographical regions, for example comparative drift distance which is likely to be influenced by the flow regime.

Drift Distance

	Met. Police 2004-2015 All Cases	Met. Police 2004-2015 Western Area Cases	Thurston 1956-59 All Cases
under 1/4 mile (0.4km)	6	3	20
1/4 to 1/2 mile (0.4 - 0.8km)	4	3	5
1/2 to 5 miles (0.8 - 8 km)	21	6	5
5 miles (8.0 km)	0	0	1
6 miles (9.7km)	1	0	1
7 miles and over (>11.3km)	1	0	1
	33	12	33

Table 2: Drift distance of bodies recovered in the two datasets

Table 2 shows the drift distance of the recovered bodies following the format and data bins adopted by Thurston (1). The notable differences are that for the 1956-59 cases:

'Mostly, the bodies were found very near to the point where they entered the water' (p 196), with 20 out of 33 cases (61%) found within 0.4 km as opposed to 3 out of 12 cases (25%) for 2004-2015. In the present study 6 out of 12 cases (50%) were found between 0.8 and 8 km away against 5 out of 20 cases (25%) in 1956-59.

There were 3 out of 33 (9%) of subjects found at 5,6 and 7 miles that were categorised as having all entered the water above Richmond (1), and there were no similar cases in 2004-2015.

No instances of long-travelled cases over 5 miles (around 8 km) were recorded in the Western Area in the present study. A possible reason is that there were no cases of bodies being recovered in the MPU dataset with both entry and exit points above Richmond rather than a fundamental change to the flow regime or other external factors. Therefore no direct comparison can be made for this reach.

Whilst the "Western Area" corresponds most closely to the area covered by Thurston's study, the number of cases for 2004-15 for this area alone is insufficient for detailed statistical analysis. However, the total number of cases in the present study is identical to that for the period included in Thurston (1956-59). By grouping both datasets into just three distance categories (< 0.5 miles, 0.5 - 5 miles and > 5 miles), it was possible to compare the distributions of drift distances for the two datasets using a χ^2 test, indicating a very highly significant difference in distribution ($\chi^2 = 16.47$, df = 2, $p < 0.0003$). The modal drift distance category for the Thurston data was for bodies which had drifted less than 0.25 mile, whilst for the 2004-15 data, the modal group was for drift distances between 0.5 and 5 miles. However, the numbers of bodies found having undergone very large drift distances (> 5 miles) were very similar for both datasets (Table 2). This implies that on average, excluding the longer drift distances, drift distances have increased since Thurston's paper. However, since Thurston's data is only available in grouped form, it is not possible to perform a more quantitative statistical test (such as a t-test or Mann-Whitney test) on the two sets of drift distances.

Factors that may have contributed to the observed changes in intervening times include: flow-regimes in the Thames, predominant clothing fashions and fabrics, and changes to likely time to discovery. Table 3 gives a qualitative analysis of the expected influence of trends on factors responsible for body drift. The apparent decrease in the proportion of bodies found close to their entry points implies that the post-1960 changes should indicate qualitative factors with an increasing expected influence on typical drift distances, however, the opposite effect is observed in the drowning statistics presented in this paper.

Drift Factor	Change from 1960s	Expected Influence on Body Drift
Flow Regime	Decline in peak flood events (3) Non-significant annual mean flow at Kingston Gauging Station (1945-2018)	Negative No change
Clothing	Lighter Modern Fabrics (35, pp.21-43, 36 pp.948-956)	Negative
Time to Discovery	Increased monitoring of River reaches (6)	Negative
Tidal Flows	Thames Barrier (1982) Sea level rise	Only during extreme events such as exceptional high tides or storm surges Increased frequency of exceptional events (Haigh et al 2016). Possibly ameliorated by Thames Barrier

Table 3 - Qualitative analysis of changes to factors influencing drift distance since the 1960s

Thurston attributes the reason for long travelled upstream cases being a less-marked tidal effect in the higher reaches where downward river flow predominates 'enhanced by periodic opening of sluice gates' a practice that still occurs for about 2 hours either side of high water. There was a major refurbishment of the Richmond Lock and Weir in the early 1990s which may have affected the flow regime (Port of London Authority, 2020)

Sex

Of the 82 cases recorded in Thurston's paper there was less than one female case to three male cases, which contrasted with the overall suicide case data in England and Wales (1956-59) of two female to three male suicides (1). Table 4. Shows the comparative break-down of the cases by sex between two datasets, the MPU data showing an even lower female to male ratio of one to thirty-two.

	2004-2015 All Cases	**Thurston Cases**
Male	32 [97%]	64 [78%]
Female	1 [3%]	18 [22%]
	33	**82**

Table 4: Sex of bodies recovered in the two datasets

Data is available from the UK Office for National Statistics (ONS) on numbers, by sex, for both all suicides in the UK, and for suicides by drowning in the UK, year by year (Office for National Statistics. Statistical Bulletin:, 2017). Comparison of both the Thurston and the Met. se) Police 2004-15 data with the ONS data for 2016 indicate significant differences ($p < 0.02$) in each case between the sex ratios for each of the two datasets, using either a χ^2 test or an "exact" test based on a binomial distribution. A particularly high significance was obtained, as expected from visual inspection of the datasets, between the sex ratios for the Thurston and the Met. Police 2004-15 data ($p < 0.003$), indicating that the observed differences in male/female suicide rates between these two studies is not due to chance.

Both Thurston's dataset and the Met. Police 2004-15 dataset indicate a significant bias towards male subjects as victims of suicides by drowning in the Thames which is in marked contrast to the European data regarding deaths by drowning attributable to suicide (Lunetta and Connolly, 2014). This male-female sex bias increases significantly between the 1950s and 1960s and the 2000s where in the latter only 1 of the 33 cases is female.

Seasonality & Phases of the Moon

One of the patterns observed by Thurston was that there was an "increase in drownings in early autumn" (1 p197) (Fig 4). Partonen et. al. (2004) observed seasonality influences on Finnish suicides which were potentially attributable to changes in luminance and temperature, that could affect intrinsic time-keeping mechanisms that influence individual body-clocks. The comparative seasonality of the drowning cases between the Thurston and contemporary dataset is shown by month (Table 5a) and by meteorological season (Table 5b). In contrast to the Thurston data, the Met. Police 2004-15 data shows a marked *decrease* during the Winter months with only 2 out of the 33 cases being between December and February.

	Met. Police 2004-2015 (All Cases)	Thurston 1956-59 (All Cases)
January	0	6
February	1	3
March	4	6
April	4	5
May	3	6
June	2	9
July	1	3
August	6	4
September	3	8
October	6	12
November	2	6
December	1	6
Total	**33**	**74**

Table 5a: Variation of bodies recovered in the datasets by month

	Met. Police 2004-2015 All Cases	Thurston 1956-59 All Cases
Spring (March - May)	11	17
Summer (June - August)	9	16
Autumn (Sept - Nov)	11	26
Winter (Dec - Feb)	2	15
Total	33	74

Table 5b: Variation of bodies recovered in the datasets by meteorological season

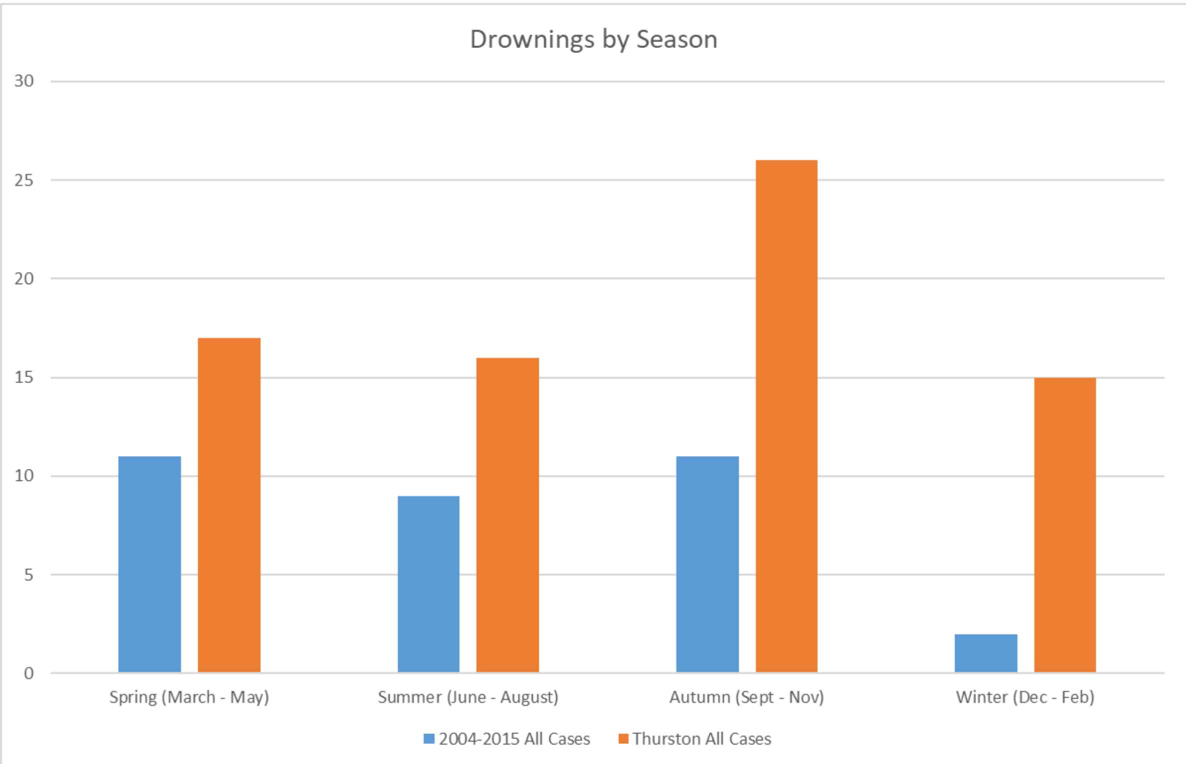


Figure 4: Comparative Seasonal Variations in Reported Drownings: 1956-59 versus 2004-2015

Using the data on the 74 cases for which the month of the case was recorded in Thurston's dataset, and the complete dataset of Met Police 2004-15 cases, a collection of Chi Squared tests on the seasonal data (Meteorological Spring, Summer, Autumn, Winter respectively) were carried out, both within each dataset and between the datasets. None of these proved statistically significant at the 5% level of significance, although in both individual datasets one season did appear to have a very different number of cases (Winter had fewer cases in the Met 2004-15 data, whereas Autumn had more cases for Thurston's data). The test for differences in number of cases between seasons for the Met Police 2004-15 data test was close to significance ($p \approx 0.084$), but those for differences between seasons within Thurston's data gave $p \approx 0.24$ and were not significant, and similarly comparing numbers by season between the two datasets ($p \approx 0.234$), did not give significance.

Variation in Cases by Lunar Phase

There is no data from Thurston the relationship between cases and the phase of the lunar cycle. However there is a tradition from antiquity which associates the full moon with affective disorders (Owens and McGowan, 2006) and the etymology of the terms lunacy and lunatic have the root *luna*- derived from the latin for moon. As discussed in Greatbatch et al (2019) there was no indication of any statistically significant increase in a variety of emergency incidents during full moon periods, however there is some divergence in the results of studies, for example Partonen et al. (2004) noted the marked increases around both Full Moon and New Moon phases, which could influence luminance levels although the impact of this in a relatively well-lit metropolitan area such as London is likely to be diminished compared to a study over a sparsely populated country like Finland.

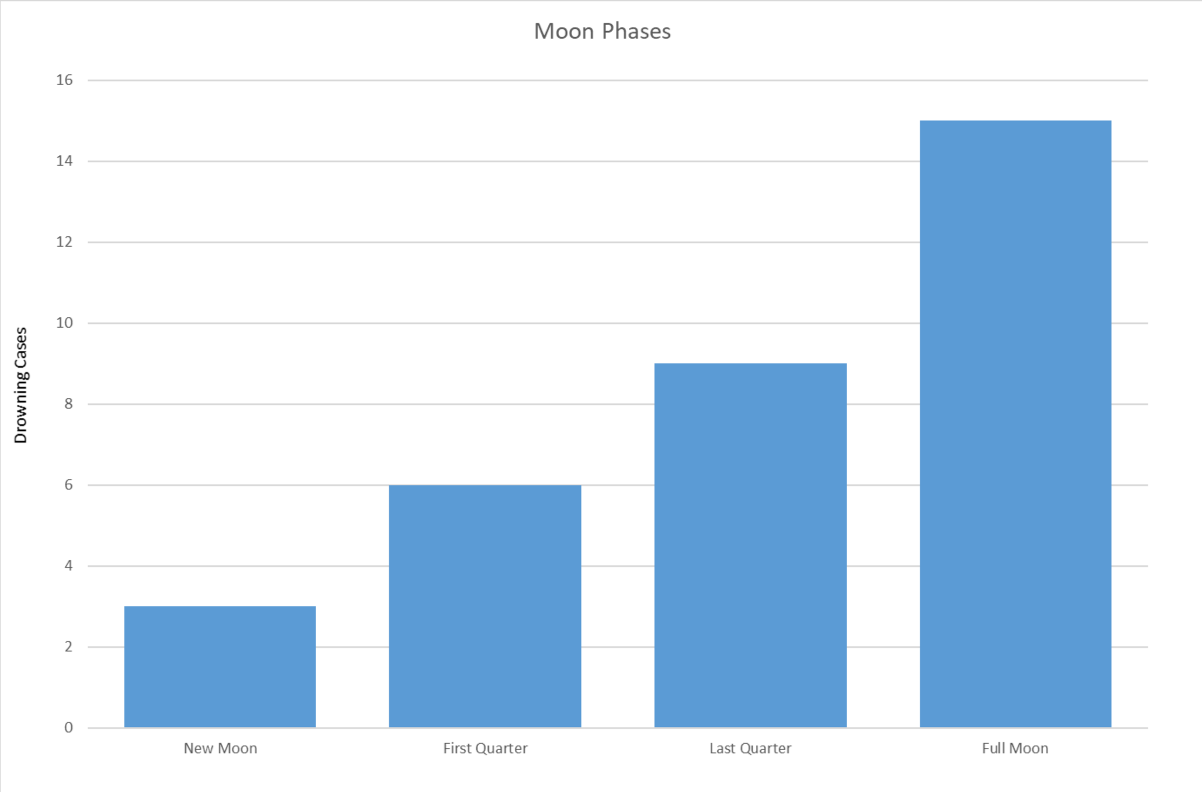


Figure 5: Lunar Phases for reported Drownings in the River Thames 2004-2015 (Metropolitan Police)

Statistical tests were carried out to investigate the significance of the variation of cases in the new (Met Police 2004-15) dataset by phase of the lunar cycle. Considering all four "quarter phases" of the lunar cycle, a Chi-Squared test gave a statistically significant result ($\chi^2 = 9.55$, DF = 3, $p \approx 0.023$), and comparing the "Full" quarter against the other three gave highly significant results both using a Chi-Squared ($\chi^2 = 7.36$, DF = 1, $p \approx 0.0067$) and a binomial "exact" test ($p \approx 0.0084$). This indicates that significantly more drowning cases occur (with the Met Police 2004-15 dataset) during the "full" quarter of the lunar cycle than at any other point of the lunar cycle.

There was a broad range of PMSIs with most being in the water under 5 days, but with 8 cases being submerged for between 10 and 20 days (Table 2, Figure 3). The range of distances travelled after entering the water was from 17m to 14.45km, with an average drift distance of 3.27km and 9 victims travelling more than 5km after entry. Thurston suggested an increase in 'early autumn' and August and October do represent the busiest months in this data, with no incidents in January (although this data only represents those victims with a known entry point, so there may well be data for those periods in the overall body recovery records). The distribution of cases by meteorological season can be seen in Table 8, and the spring and autumn rises in incidents observed. The average PMSI for winter is higher than any other season, which could suggest an impact of cold water on decomposition, but with only two cases in that season it is impossible to hypothesise with any confidence. Friday and Saturday were the most common days for the deaths, with Sunday the least common.

Clothing, Distance Travelled ("Downdrift") and Time in Water

There are 26 records in the present study that contain information on clothing, and these have been subdivided into groups by clothing weight. These weights were compared to distance travelled using Spearman's rank correlation coefficient since the clothing weights followed an ordinal rather than continuous scale. The results (Table 7) showed that there were weak (not statistically significant at 5% level) correlations between clothing weight time spent in the water, with heavier clothing tending to imply that the body spent less time in the water. Correlations between clothing weight and distance travelled (downdrift) and between time in the water and distance travelled were so weak as to be considered negligible ($p > 0.75$ in both cases).

The negligible correlation between time in water and net distance travelled between point of entry and point of removal from the river can be observed in the scatterplot (Figure 6). Negative distances (negative downdrifts) indicate that the particular body was displaced upstream, rather than the expected downstream. There is clearly no simple relationship between net distance travelled and time in the water. This is most probably again due to the oscillatory nature of the effect of tides, since all bodies had been in the river for seven days or more.

		days in water	Clothing score	downdrift
Days in Water	Correlation Coefficient	1.00	-0.18	0.06
	Sig. (2-tailed)		0.32	0.76
	N	33	33	33
Clothing Score	Correlation Coefficient	-0.18	1.00	0.00
	Sig. (2-tailed)	0.32		0.99
	N	33	33	33
Downdrift	Correlation Coefficient	0.06	0.00	1.00
	Sig. (2-tailed)	0.76	0.99	
	N	33	33	33

Table 6: Correlation results for clothing weight versus distance travelled (downdrift) and time in water. Spearman's Rank Correlation Coefficient was used with a two-tailed test of statistical significance. None of the correlations were statistically significant at the 5% level.

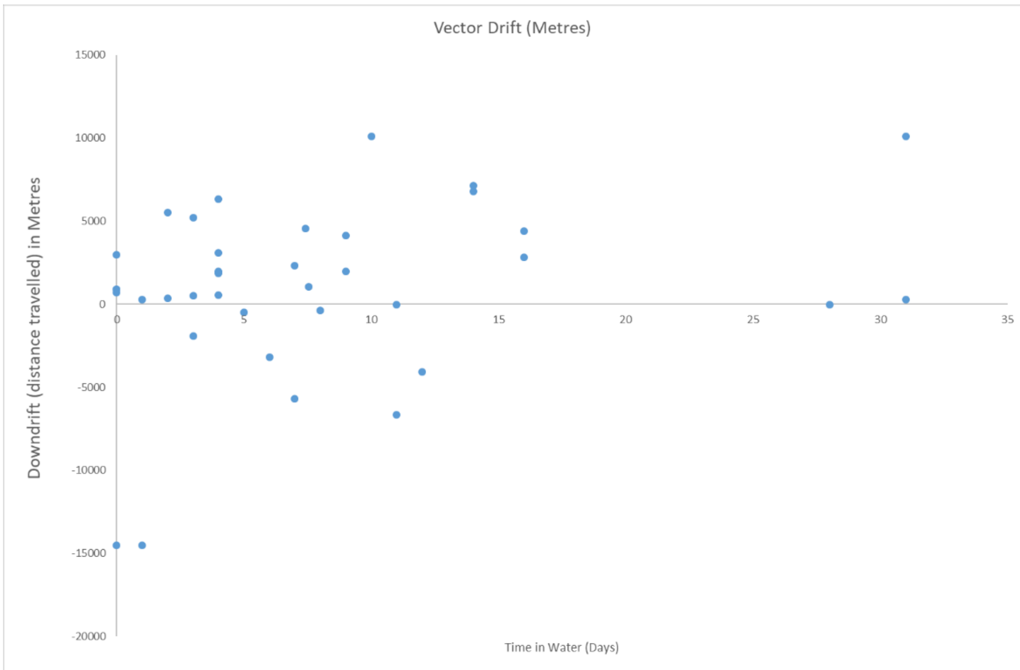


Figure 6: Downdrift of Recovered Bodies (metres) against Time in Water (days) 2004-2015 (Metropolitan Police)

Discussion & Conclusion

Clearly this an unusual dataset, and one that could only have been created or collected by an agency such as the River Police. It deals with a subset of recoveries from the river - those where a point of entry is known - and as such is prone to some of the well understood limitations of small datasets. Any pattern or inference made from the data must be considered in terms of the overall set size, yet in order to calculate the drift or submersion time of bodies, this is the only reliable data that could be used. Future technological advances, such as wearable technology, or an increase in the CCTV coverage on the River Thames may result in more complete datasets in the future.

The analysis has produced some notable results, which may have policy implications. Firstly, it can be seen that there are two seasonal spikes - around spring and autumn. For this we have two working hypotheses:-

1. It is related to a seasonal disorder: The daylight is either decreasing, summer holidays are finished, and the winter is coming. In late spring the later, lighter nights may reinforce feelings of isolation, if lonely people can observe activities occurring outside, yet they are not involved.
2. It is physically more dangerous: The two spikes coincide roughly with the equinoxes and as such the river is at its highest astronomical tidal heights. The driving factor behind an increase in deaths may be that the chances of survival are diminished at these times, rather than the input number of attempted suicides.

The negative correlation between clothing weight and time spent in water is perhaps counterintuitive. However, a possible explanation for this could be that the clothing level recorded is that observed when each body was recovered from the water, but not at the time of the body entering the river. The longer the body remained in the water, particularly bearing in mind the oscillatory influence of tides, the greater the opportunity for clothing to be removed from the body by natural processes such as abrasion against the riverbed, riverbanks or other objects in the water, as suggested by Armstrong & Erskine (13).

There is a dramatic imbalance of men to women, which appears to run counter to global, or regional statistics, and this may be something addressable through education, or social policy.

Many of the current campaigns such as the Royal Society for the Prevention of Accidents (ROSPA) “don’t drink and drown” or the seasonal work concerning Men Missing on a Night Out (Newiss and Greatbatch, 2019) specifically address male accidental drowning and suicide around the Christmas holidays. This work potentially adds insight to male suicide figures in London, in terms of both seasonality and spatiality, which in turn may inform policy.

This work serves as a baseline study, in a field where little work has been completed (with the notable exception of Dennison-Wilkins 2021) and provides a set of analysed data and conclusions which would be hard to replicate without the input of the MPU. This work has the potential to form the basis of heuristics for body recovery search planning, and to inform search patterns for drowned victims in other bodies of water. The importance of this should not be underestimated, when considering the emotional impact on the families of drowned persons; in essence any tool that reduces the PMSI is likely to reduce traumatic impact on survivors.

As this work is essentially a baseline study with a specialist dataset, and limited to that, it provides much scope for further work. Increasing the size of the dataset would allow for a greater range of statistical tests to be performed, but would also give more confidence in some of the observed patterns. The size of the dataset could be improved by recording more cases of recorded entry into the water - by increasing the monitoring and recording on Thames - and matching them up with subsequent body recoveries. Other rivers could also be incorporated, and a comparative study of rivers performed. There is also potential for retrospective work on recovery-only cases, applying CCTV and other information from police or coroners reports to determine the point of entry.

Further experimental work could also be performed with clothing, using fixed dummies, or private rivers to investigate the rate of deterioration of clothing, and potentially even the impacts of different materials and garments on the movement and buoyancy of human bodies. Finally, there is some existing application to forensic science and policing in the analytical results concerning drift, but there is some potential in future work to use some of the results here to generate heuristics for predicting body location.

(no. of cases)	Gender		Age			Date of Entry (Most Frequent)			Drift Direction	
	Male	Female	Ave.	Max	Min	Season	Month	Moon Phase	Downstream	Upstream
All	32	1	36	79	20	{Spring, Autumn} n=11	{August, October} n=6	{Full Moon} n=15	23	10
Western Reach	12	0	39	62	21	{Summer} n=5	{August} n=3	{Quarter Moon} n=5	8	4
Central/Eastern	20	1	34	79	20	{Spring} n=9	{April, October} n=4	{Full Moon} n=11	15	6

	Drift Distance (m)			PMSI (Days)		
	Average	Max	Min	Average	Max	Min
All	3,273	14,517	16	8	31	0
Downstream	3,089	10,102	284	7	31	0
Upstream	3,696	14,517	16	9	28	1

Sex	Age	Drift Direction	Drift Distance (m)	PMSI (days)	Clothing weight	Month (Entry)	Moon Phase (Entry)	Season (Meteorological)
Male	35	UPSTREAM	4,089	12	3	MARCH	Last Quarter	SPRING
Male	20	DOWNSTREAM	4,408	16	1	MARCH	Full Moon	SPRING
Male	34	DOWNSTREAM	2,805	16	2	MARCH	New Moon	SPRING
Male	39	UPSTREAM	6,654	11	3	MARCH	First Quarter	SPRING
Male	23	UPSTREAM	377	8	2	APRIL	First Quarter	SPRING
Male	28	DOWNSTREAM	4,111	9	0	APRIL	Last Quarter	SPRING
Male	29	DOWNSTREAM	1,958	9	3	APRIL	Last Quarter	SPRING
Male	28	DOWNSTREAM	2,960	0	3	APRIL	Full Moon	SPRING
Male	59	DOWNSTREAM	2,336	7	0	MAY	Full Moon	SPRING
Male	45	UPSTREAM	3,192	6	0	MAY	Full Moon	SPRING
Male	30	UPSTREAM	475	5	2	MAY	Last Quarter	SPRING
Male	37	DOWNSTREAM	6,310	4	1	JUNE	New Moon	SUMMER
Male	35	DOWNSTREAM	291	31	4	JUNE	Full Moon	SUMMER
Male	54	DOWNSTREAM	897	0	3	JULY	Full Moon	SUMMER
Male	29	DOWNSTREAM	368	2	0	AUGUST	Full Moon	SUMMER
Male	40	UPSTREAM	1,931	3	4	AUGUST	Full Moon	SUMMER
Male	31	DOWNSTREAM	1,961	4	3	AUGUST	Full Moon	SUMMER
Male	45	DOWNSTREAM	284	1	4	AUGUST	Last Quarter	SUMMER
Male	21	DOWNSTREAM	3,084	4	0	AUGUST	Last Quarter	SUMMER
Male	27	DOWNSTREAM	1,869	4	2	AUGUST	Last Quarter	SUMMER
Male	24	DOWNSTREAM	5,212	3	2	SEPTEMBER	New Moon	AUTUMN
Male	23	DOWNSTREAM	554	4	2	SEPTEMBER	Last Quarter	AUTUMN
Male	32	DOWNSTREAM	5,501	2	3	SEPTEMBER	Full Moon	AUTUMN
Male	29	UPSTREAM	25	11	3	OCTOBER	Full Moon	AUTUMN
Male	62	UPSTREAM	5,687	7	1	OCTOBER	First Quarter	AUTUMN
Male	25	DOWNSTREAM	10,102	10	3	OCTOBER	First Quarter	AUTUMN
Male	50	DOWNSTREAM	525	3	3	OCTOBER	Full Moon	AUTUMN
Male	23	DOWNSTREAM	7,127	14	2	OCTOBER	Last Quarter	AUTUMN
Male	45	DOWNSTREAM	901	0	2	OCTOBER	Full Moon	AUTUMN
Female	79	UPSTREAM	14,517	1	3	NOVEMBER	Full Moon	AUTUMN
Male	25	DOWNSTREAM	6,768	14	0	NOVEMBER	First Quarter	AUTUMN
Male	48	DOWNSTREAM	709	0	0	FEBRUARY	Full Moon	WINTER
Male	23	UPSTREAM	16	28	4	DECEMBER	First Quarter	WINTER

Table 9. All data, anonymised.

About the authors

Dr Ian Greatbatch is an Officer in the UK Fire Service, Visiting Researcher at the University of Portsmouth and formerly Associate Professor at Kingston University London. His research concerns the operational effectiveness of Search and Rescue and Fire and Rescue Operations.

David Livingstone currently lectures in Computer Science at Kingston University where he was a senior lecturer from 1994 until retirement in 2017. He is also a Visiting Research Fellow in the Faculty of Science Engineering and Computing, Kingston University where his major research areas are GIS and Environmental Modelling applied to Coastal Geomorphology and Innovative Learning Environments for Computer Programming.

Dr Gordon Hunter is Associate Professor in Mathematics and Computer Science at Kingston University, where he has worked since 2003. Originally a Mathematical Physicist, his recent research interests include various applications of Statistics and Machine Learning in the Environmental and Medical/Life Sciences.

Yvonne Saunderson, is a former Police Sergeant on the Metropolitan Police Marine Policing Unit. As part of the Marine Policing Unit, Yvonne had responsibility for coordinating the police search and rescue response on the River Thames in London. Part of this work involved searching for missing persons who were believed to have drowned in the river, the recovery of bodies from the water and assisting with the follow up investigation and identification of the individuals concerned.

References

Aguglia, A., Serafini, G., Solano, P., Giacomini, G., Conigliaro, C., Salvi, V., Mencacci, C., *et al.* (2019), “The role of seasonality and photoperiod on the lethality of suicide attempts: A case-control study”, *Journal of Affective Disorders*, Vol. 246, pp. 895–901.

Armstrong, E.J. and Erskine, K.L. (2011), *Water-Related Death Investigation : Practical Methods and Forensic Applications*, CRC Press, Boca Raton, FL.

Branche, C. and van Beeck, E. (2014), “Summary and Recommendations”, in Bierens, J.J.L.M. (Ed.), *Drowning: Prevention, Rescue, Treatment*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 81–83.

Butcher, L. (2010), “Shipping: safety on the River Thames and the Marchioness disaster”, House of Commons Library, available at: https://researchbriefings.parliament.uk/ResearchBriefing/Summary/SN00769 (accessed 15 December 2019).

Byard, R.W. (2018), “Putrefaction: An Additional Complicating Factor in the Assessment of Freshwater Drownings in Rivers”, *Journal of Forensic Sciences*, Vol. 63 No. 3, pp. 899–901.

Byard, R.W., Houldsworth, G., James, R.A. and Gilbert, J.D. (2001), “Characteristic Features of Suicidal Drownings: A 20-Year Study”, *The American Journal of Forensic Medicine and Pathology*, Vol. 22 No. 2, pp. 134–138.

Chiou, C.-R., Tsai, W.-L. and Leung, Y.-F. (2010), “A GIS-dynamic segmentation approach to planning travel routes on forest trail networks in Central Taiwan”, *Landscape and Urban Planning*, Vol. 97 No. 4, pp. 221–228.

Copeland, A.R. (1987), “Suicide by Drowning”:, *The American Journal of Forensic Medicine and Pathology*, Vol. 8 No. 1, pp. 18–22.

Crooks, S.M. (1994), “CHANGING FLOOD PEAK LEVELS ON THE RIVER THAMES.”, *Proceedings of the Institution of Civil Engineers - Water Maritime and Energy*, Vol. 106 No. 3, pp. 267–279.

van Daalen, M.A., de Kat, D.S., Oude Grotebevelsborg, B.F.L., de Leeuwe, R., Warnaar, J., Oostra, R.J. and M Duijst-Heesters, W.L.J. (2017), “An Aquatic Decomposition Scoring Method to Potentially Predict the Postmortem Submersion Interval of Bodies Recovered from the North Sea”, *Journal of Forensic Sciences*, Vol. 62 No. 2, pp. 369–373.

Dennison-Wilkins, L. (2021) A Study of Human Body Movement in Inland Waterways. [Doctoral dissertation, University of Dundee, Leverhulme Research Centre for Forensic Science]

Greatbatch, I., Koester, R.J. and Kleinsmith, A.L. (2019), “Rescue responses during a full moon and Friday 13th”, *International Journal of Emergency Services*, Vol. 8 No. 2, pp. 108–121.

Haw, C. and Hawton, K. (2016), “Suicide and Self-Harm by Drowning: A Review of the Literature”, *Archives of Suicide Research*, Vol. 20 No. 2, pp. 95–112.

Haywood, I. (1998), “London”, *Cities*, Vol. 15 No. 5, pp. 381–392.

Heaton, V., Lagden, A., Moffatt, C. and Simmons, T. (2010), “Predicting the Postmortem Submersion Interval for Human Remains Recovered from U.K. Waterways”, *Journal of Forensic Sciences*, Vol. 55 No. 2, pp. 302–307.

Helmus, J., Zorell, S., Bajanowski, T. and Poetsch, M. (2018), “Persistence of DNA on clothes after exposure to water for different time periods—a study on bathtub, pond, and river”, *International Journal of Legal Medicine*, Vol. 132 No. 1, pp. 99–106.

Humphreys, M.K., Panacek, E., Green, W. and Albers, E. (2013), “Comparison of Protocols for Measuring and Calculating Postmortem Submersion Intervals for Human Analogs in Fresh Water”, *Journal of Forensic Sciences*, Vol. 58 No. 2, pp. 513–517.

Lasota, D., Pawłowski, W., Krajewski, P., Staniszewska, A., Goniewicz, K. and Goniewicz, M. (2019), “Seasonality of Suicides among Victims Who Are under the Influence of Alcohol”, *International Journal of Environmental Research and Public Health*, Vol. 16 No. 15, p. 2806.

Lucas, J., Goldfeder, L.B. and Gill, J.R. (2002), “Bodies found in the waterways of New York City.”, *Journal of Forensic Sciences*, Vol. 47 No. 1, pp. 137–141.

Lukmanji, A., Williams, J.V.A., Bulloch, A.G.M., Bhattarai, A. and Patten, S.B. (2019), “Seasonal variation in symptoms of depression: A Canadian population based study”, *Journal of Affective Disorders*, Vol. 255, pp. 142–149.

Lunetta, P. and Connolly, J. (2014), “Risk Factors for Drowning: Suicidal Drowning”, in Bierens, J.J.L.M. (Ed.), *Drowning*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 131–134.

Lunetta, P., Zaferes, A. and Modell, J. (2014), “Establishing the Cause and Manner of Death for Bodies Found in Water”, in Bierens, J.J.L.M. (Ed.), *Drowning*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1179–1189.

Martyn, M. (2014), “The Global Burden of Drowning”, in Bierens, J.J.L.M. (Ed.), *Drowning*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 91–94.

Mateus, M., de Pablo, H. and Vaz, N. (2013), “An investigation on body displacement after two drowning accidents”, *Forensic Science International*, Vol. 229 No. 1–3, pp. e6–e12.

Newiss, G. and Greatbatch, I. (2019), “The spatiality of men who go missing on a night out: Implications for risk assessment and search strategies”, *International Journal of Emergency Services*, Vol. ahead-of-print No. ahead-of-print, available at:https://doi.org/10.1108/IJES-03-2019-0012.

Office for National Statistics. Statistical Bulletin: (2017), *Suicides in the UK: 2016 Registrations.*

Owens, M. and McGowan, I. (2006), “Madness and the moon: The lunar cycle and psychopathology”, *German Journal of Psychiatry*, Vol. 9.

Partonen, T., Haukka, J., Pirkola, S., Isometsa, E. and Lonnqvist, J. (2004), “Time patterns and seasonal mismatch in suicide”, *Acta Psychiatrica Scandinavica*, Vol. 109 No. 2, pp. 110–115.

Port of London Authority. (2020), “Richmond Lock and Weir”, available at: http://www.pla.co.uk/About-Us/Richmond-Lock-and-Weir (accessed 1 November 2020).

Reijnen, G., Gelderman, H.T., Oude Grotebevelsborg, B.F.L., Reijnders, U.J.L. and Duijst, W.L.J.M. (2018), “The correlation between the Aquatic Decomposition Score (ADS) and the post-mortem submersion interval measured in Accumulated Degree Days (ADD) in bodies recovered from fresh water”, *Forensic Science, Medicine and Pathology*, Vol. 14 No. 3, pp. 301–306.

ROSPA. (2007), “Royal Society for the Prevention of Accidents: Child holiday swimming pool safety”, available at: https://www.rospa.com/Leisure-Safety/Water/Advice/Holiday-Swimming-Pool (accessed 17 December 2019).

Rotton, J. and Kelly, I.W. (1985), “Much ado about the full moon: A meta-analysis of lunar-lunacy research.”, *Psychological Bulletin*, Vol. 97 No. 2, pp. 286–306.

Saxena, S., Krug, E.G., Chestnov, O. and World Health Organization (Eds.). (2014), *Preventing Suicide: A Global Imperative*, World Health Organization, Geneva.

Skogan, W.G. (2019), “The future of CCTV”, *Criminology & Public Policy*, Vol. 18 No. 1, pp. 161–166.

UK Met. Office. (2020), "Kew Gardens Climate Station Data Page", available at: https://www.metoffice.gov.uk/research/climate/maps-and-data/uk-climate-averages/gcpuckhb6 (accessed 1 October 2020).

Thurston, G. (1960), "Bodies in the West London Thames", Medico-Legal Journal, Vol.28, pp. 193-199.

Varnik, A., Kolves, K., van der Feltz-Cornelis, C.M., Marusic, A., Oskarsson, H., Palmer, A., Reisch, T., *et al.* (2008), "Suicide methods in Europe: a gender-specific analysis of countries participating in the 'European Alliance Against Depression'", *Journal of Epidemiology & Community Health*, Vol. 62 No. 6, pp. 545–551.

Ward, R.C. and Robinson, M. (2000), *Principles of Hydrology*, 4. ed., McGraw-Hill, London.

Werner, A. (2015), "Nineteenth- and Twentieth-Century Photographs of the River Thames: The Collections of the Museum of London and the Port of London Authority", *The London Journal*, Vol. 40 No. 3, pp. 196–210.

Williams, A., Rogers, C.J. and Cassella, J.P. (2019), "Why does the UK need a Human Taphonomy Facility?", *Forensic Science International*, Vol. 296, pp. 74–79.