Designing a detection experiment: tricks and trade-offs

Cindy E. Hauser[1], Joslin L. Moore[2], Katherine M. Giljohann[1], Georgia E. Garrard[3] and Michael A. McCarthy[3]
[1] Department of Resource Management and Geography, University of Melbourne – Burnley Campus, 500 Yarra Blvd, Richmond 3121
[2] Australian Research Centre for Urban Ecology, Royal Botanic Gardens Melbourne c/o University of Melbourne, Parkville 3010
[3] School of Botany, University of Melbourne, Parkville 3010
(chauser@unimelb.edu.au)

Summary Detection experiments are relatively new to weed management and provide an opportunity to assess the capability of weed surveys. Long-held statistical principles such as stratification, randomisation and replication should guide detection experiment design, but these can conflict with the need to create a realistic search environment, and the logistics of conducting the experiment. In this paper we outline some of these principles and trade-offs, with particular reference to a recent experiment on the detectability of hawkweeds in the Victorian Alps.

Keywords Eradication, surveys, experimental design, *Hieracium* species.

INTRODUCTION

Weed eradication and management requires successful detection before control can take place. Yet even the best survey methods and designs harbour some possibility that individual weeds or entire infestations may be overlooked. Imperfect weed detection should be accounted for when interpreting survey data and choosing actions. This allows us to manage the risks of failing to recognise a new incursion (Mehta *et al.* 2007), continuing weed impacts (Hauser and McCarthy 2009), weed spread (Emry *et al.* 2001) and/or erroneously declaring eradication (Regan *et al.* 2006, Rout *et al.* 2009).

While methods exist to incorporate imperfect detection into weed management plans, most rely on expert opinion to parameterise detection rates. Only a few recent studies have directly estimated detection rates for plant surveys (Garrard *et al.* 2008, Chen *et al.* 2009, Moore *et al.* 2011). Controlled experiments can estimate detection rates directly and identify the factors that influence successful detection. Target plant characteristics, the surrounding environment and searchers' attributes can be varied and replicated for optimal statistical inference. Weed management plans can subsequently be adjusted to address the conditions at hand.

We have designed, conducted and analysed a number of detection experiments in the contexts of weed and threatened species management. In this paper we will discuss the design of detection experiments, with particular reference to a recent experiment using hawkweeds in the Victorian Alps.

HAWKWEED MANAGEMENT IN ALPINE VICTORIA

Hieracium species are recognised as a serious threat to agriculture and the environment in Australia (Williams and Holland 2007). Orange hawkweed (*Hieracium aurantiacum*) and king devil hawkweed (*Hieracium praealtum*) were introduced to Falls Creek, Victoria and neighbouring ski runs. These incursions have naturalised and spread into the adjoining Alpine National Park, were discovered in 1999 and 2003 respectively, and are the subjects of an eradication program.

Hawkweeds lie dormant under snow over winter and then emerge in spring, flowering in summer. Surveys to detect new hawkweed infestations are conducted in the late spring and summer months when the plants are most conspicuous, with search teams walking parallel transects that cover the search area. Individual infestations can be sparsely distributed, so teams may search for days without detections.

University of Melbourne researchers have helped managing agencies identify areas at highest risk of infestation. Models incorporating seed dispersal by wind and habitat suitability predict where seeds are likely to arrive and successfully germinate, producing a map of likely occurrence (Williams *et al.* 2008). This map can then be combined with a detection model to determine how much search effort should be allocated to each location in the landscape (Hauser and McCarthy 2009).

Parameters for the detection model were initially estimated by experts, and a subsequent study used search simulations in the field to directly estimate the relationship between search effort, measured in hours per hectare, and the probability of detecting non-flowering orange hawkweed (Moore *et al.* 2011). However the different hawkweed forms are also expected to influence detectability. Orange hawkweed flowers are distinct from all other species occurring in the Alpine National Park; king devil hawkweed's yellow flowers may set it apart from the dominating grasses and heath but they look similar to other locally occurring species such as *Hypochaeris radicata*. We designed a subsequent experiment to estimate the relative detectability of these different species and forms against a variety of background vegetation.

EXPERIMENTAL DESIGN

Strong statistical inferences can only be made when an experiment is well designed. It is important to identify important influencing variables as part of the design; these will inform the site(s) to be selected, the targets to be used, the targets' placement within sites and the role of searchers within the experiment.

Identifying variables It is important to develop *a priori* hypotheses of what factors influence plant detection. We considered factors that relate to the target plant characteristics, the surrounding environment and the searchers' experience, and sought to measure or control these in the experiment.

In plant surveys, search effort may be adjusted to improve survey effectiveness or efficiency. We aimed to build a statistical model linking the time taken to detect a target hawkweed to influencing factors. Potential influencing factors included:

1. *Target plant characteristics:* species (orange vs King devil), form (flowering vs non-flowering rosette), number of flowers if present.
2. *The surrounding environment:* dominant background vegetation (grass vs heath), density of other yellow-flowered species, weather.
3. *Searcher attributes:* training and knowledge, use of a hat or sunglasses, colour blindness, gender (to account for possible undiagnosed colour-blindness), and fatigue.

In our experiment, target plant characteristics, dominant background vegetation and density of yellow flowers were controlled, while the other factors were measured but not controlled.

Site selection Experimental sites should reflect the real search environment as much as possible, though some features may need to be controlled and replicated across plots. To allow for replication, individual plots may be smaller than the area typically surveyed for weeds. However plot size must allow for and even encourage search behaviour that resembles real surveys. Some of the factors influencing detection occur at the plot level and must be assessed or measured directly. The logistics of accessing plots must also be considered.

Hawkweed management prioritisation operates at a 20 m GIS grid size (Hauser and McCarthy 2009), so we selected 20 m × 20 m plots. Since we were interested in the influence of background vegetation and yellow-flowered species, we selected plots that were relatively homogeneous with respect to these factors and replicated all combinations of the factors (see Table 1). While we aimed for three replicates of each plot combination, it was not always possible to construct three homogeneous plots within 15 minutes' walking distance.

In addition to these initial subjective assessments of background vegetation and yellow flower density, we took measurements across each plot. Vegetation height was measured at 2 m intervals across each plot. To estimate yellow flower density, we stratified each plot into 25 × (4 m)^2 grids and from each grid randomly selected a 1 m^2 quadrat for sampling (thus covering 6.25% of the plot area). We photographed each quadrat and counted each yellow flower,

Table 1. Initial assessments and target allocations for the 16 plots selected in the hawkweed detection experiment.

Plot characteristics			Target plantings				
Plot ID	Dominant vegetation	Yellow flower density	flowering orange	flowering King Devil	non-flowering orange	non-flowering King Devil	Total targets
A	grass	high	0	2	0	1	3
B	grass	high	2	1	2	2	7
C	grass	high	1	0	1	0	2
D	heath	low	0	2	1	1	4
E	heath	low	1	0	0	2	3
F	heath	low	2	1	2	0	5
G	grass	low	1	1	0	1	3
H	grass	low	2	0	2	2	6
I	grass	low	0	2	1	0	3
J	heath	high	1	2	2	1	6
K	heath	high	2	1	1	2	6
L	mixed	high	0	0	1	2	3
M	mixed	high	1	1	2	0	4
N	mixed	high	2	2	0	1	5
O	mixed	low	1	1	2	1	5
P	mixed	low	2	2	1	2	7
	Total targets		18	18	18	18	72

classifying the flowers by species and measuring the flower diameter and height. The percentage area covered by yellow flowers could later be estimated from flower diameters and counts, and/or by counting yellow pixels in digital photographs.

Target plants Individual variation in the target species, such as size, life stage or form, may influence detection. Such variables should ideally be controlled, varied and replicated.

A detection experiment may involve opportunistic use of the target species *in situ*, translocation of individuals from outside the study area, cultivation of the species, or even species mimics. The risks associated with using a weedy species, and the possibility of escape, should be very carefully assessed. A clear protocol for the storage, placement, recovery and disposal of target plants may be necessary.

For each target individual, variables thought to influence detection should be measured, and the targets themselves should adequately mimic the species' natural placement and appearance.

We anticipated that hawkweed species and flower presence were likely to influence detection rates and therefore we controlled the quantity and placement of each target form in the experiment. To mitigate risk, we cultivated small rosettes of these State prohibited species in locked glasshouses and environmental growth chambers under permit. Due to their identical growing conditions, the rosettes were highly uniform in size and appearance. It was difficult to maintain the hawkweed rosettes in conditions comparable to alpine Victoria, and their appearance differed somewhat to hawkweed rosettes discovered in the field. The rosettes were planted at the field site in their pots, lined with an additional layer of plastic to prevent root material from contaminating the soil. The rim of each pot was covered with litter, to minimise the searchers' opportunity to alter their behaviour by seeking pot rims as a visual cue. Each target's position was recorded using a high-resolution GPS to assist with later retrieval. On completion of the experiment, a clear protocol was developed for the recovery of each individually-identified rosette that included inspection of the protecting plastic for breaches or root material. Regular revisits to the planting sites are scheduled to ensure any potential subsequent infestation is detected and controlled.

We deemed the possibility of hawkweed seed set to be unacceptable and as a consequence sought artificial mimics for hawkweed flowers. We commissioned Ewin Wood of Natural History Productions (http://www.naturalhistoryproductions.com.au/) to produce 25 orange hawkweed stems and 25 King devil stems with varying quantities of flowers per stem. The mimics bore an excellent resemblance to the species under typical survey conditions, though their artifice was clear on close inspection. We anticipated some damage to stems during the experiment due to harsh weather conditions and crushing underfoot and therefore allocated only 36 of the 50 commissioned stems to the plots, setting aside 14 stems to replace damaged stems where needed. Flowers were positioned in selected plots at a constant stem height, so that variation in flower detectability could be related to background vegetation height, as well as the colour and quantity of flowers on each stem.

Placing targets in plots Maximum statistical power can be gained by knowing the precise location of every target individual in the experiment, although methods do exist for analysing experiments where the presence, abundance and distribution of targets is unknown (Garrard *et al.* 2008). In practice, a complete census of the plots may be difficult or impossible. The alternative is to select plots where the target species is absent and control the abundance and placement of target plants.

Where individuals are distributed randomly, the detection rate should increase linearly with abundance (McCarthy *et al.* in press). The density of targets should reflect typical densities encountered during real searches; it may need to be controlled, varied and replicated across plots.

The locations of targets in plots should be randomised. Plots may need to be stratified prior to randomisation if the target species has specific microhabitat requirements. While randomisation, replication and balance are important for statistical inference, some care needs to be taken that searchers will not detect and alter their behaviour according to patterns that would not arise during a typical search.

New infestations of hawkweeds typically occur at low densities across alpine Victoria. We therefore required a low density of targets across plots, although it was not feasible to set hawkweed density as low as it generally occurs in the alps – exposing each searcher to just one or two targets per day would not provide sufficient replication to meaningfully estimate detection rates. At the other extreme, an arbitrarily large number of targets per plot would provide high replication and excellent statistical inference but interfere with the search experience – searcher behaviour would likely change as the vegetation became dominated by the target species.

We assumed that detection rates for the four hawkweed forms (flowering orange, flowering King devil, non-flowering orange, non-flowering King devil) were independent of each other. Placing an equal number of each target form in each plot would provide the best balance for statistical inference, yet searchers would be likely to detect such a pattern and alter their behaviour according to the targets they anticipate (e.g. by shifting their search image exclusively to yellow flowers after they presumed they had found 'enough' orange flowers). For this reason we randomly allocated zero, one or two plants of each target form to each plot, such that three plants of each target form were allocated to each vegetation–yellow flower stratum (see Table 1). In the absence of clear data describing hawkweeds' micro-habitat preferences in alpine Victoria, we randomly selected the location of each target within its allocated plot.

This design was not an exact simulation of real search conditions, but was identified as a reasonable trade-off between the need for rarity and the need for replication. We acknowledged that searchers' motivation might be unrealistically high as they were frequently rewarded with successful finds, and that this might alter their attentiveness and hence detection rates.

Searchers Searchers recruited to participate in a detection experiment should represent the searcher profiles involved in real weed surveys. It would be preferable to control, vary and replicate any factors that are expected to influence detection, although this may not be possible. Regardless, it is useful to measure them. Such factors typically relate to the training and experience of the searcher, as well as their physical fitness.

Studies involving human participation frequently require approval from an ethics committee. This should be sought well in advance of the experiment. Participant recruitment, health and safety, issues of privacy, anonymity and data management may need to be considered.

We submitted an application to the University human ethics committee six months in advance of the hawkweed detection experiment. This application described the purpose of the experiment, its experimental design, procedures for ethical recruitment, instruction and data management, and sample questionnaires.

We sought searcher information that we anticipated would affect detection *via* questionnaires. Questionnaires were given to participants after their searches were completed, to avoid any altered search behaviour that might arise from considering the questions posed. We asked participants about:

- their role in hawkweed searches, e.g. government employee, private contractor, volunteer;
- the amount of time they had previously spent participating in surveys for hawkweeds and other weeds;
- other forms of experience and training;
- use of a hat and/or sunglasses during searches;
- colour blindness; and
- gender.

The detection experiment was conducted in parallel to real hawkweed surveys to maximise the number of participants available. Participants were recruited in partnership with the survey managers to suit their schedules, although we negotiated as even a ratio of roles amongst the participants as practicable.

RUNNING SEARCHES

Experimental searches should ideally resemble real weed surveys as closely as possible. However, search procedures might need to be altered to provide quality data that answer the research questions. If the data typically logged during a plant survey are insufficient for the detection experiment, then additional measuring instruments may be required and new data sheets should be designed. These additional data might reasonably be entered by searchers, or pose a distraction from typical search procedures and behaviour. In the latter case researchers may need to supervise searchers and record detection data.

A draft schedule should be developed to identify what resources are available (e.g. number of searchers, number of supervisors, number of plots, time) and how they can best be allocated. It is preferable to include as many combinations of each influencing factor as possible, but some trade-offs may need to be made. For example, it may be possible to involve either a large number of searchers for a brief time each or a small number of searchers for a long time each, due to limits on time. Basic reasoning relating back to the key research questions may reveal the preferred trade-off, and computer simulations based on expected detection rates can also assist. Allocations should be randomised wherever possible, within logistical constraints, such as travel between plots.

Detection experiments address the process of information gathering and, as such, care should be taken regarding the information revealed to participants. Decisions should be based on the key research questions and the typical conditions of plant survey.

Hawkweed surveys are usually conducted in teams, with searchers forming a line and walking parallel transects to cover each designated area. However, our detection experiment was designed for individual searches. This allowed us to measure searcher-to-searcher variation and investigate the effects of training and experience. An individual-searcher approach was also far more tractable with respect to recruiting sufficient participants and constructing smaller plots. We requested that searchers walk parallel transects through plots to simulate the group's line search style.

Searchers were shown sample hawkweed rosettes and flower mimics during their introduction to the experiment. This approach is taken during real hawkweed surveys. Furthermore, we aimed to avoid confusion, disappointment or search behaviour changes that might arise from discovering a mimic for the first time during a search. However, searchers were not provided with information regarding patterns of target occurrence in the plots.

We investigated the use of GPSs carried by the searchers to store data on detection but concluded that the devices would be insufficiently accurate to identify individual targets within a plot, with data entry by searchers somewhat prone to error. The use of stopwatches and data sheets by searchers would distract substantially from the search task and also be prone to error; therefore, we elected to allocate a supervisor from the research team to each searcher. To avoid repeated detections of the same target, searchers used small flags to mark detected targets, which were removed when the search period ended.

Supervisors were not told the locations of each target, in order to prevent conscious or unconscious hints being communicated to searchers. Allocating supervisors to searchers, rather than plots, reduced the supervisors' ability to learn target locations and eased the transition of searchers between plots.

Supervisors were responsible for measuring the time taken for each target detection, and identifying each target *via* its features (e.g. species, presence of and number of flowers) and location. We were focussed on the 'active search' phase of detection and not on subsequent time spent identifying the species or otherwise managing the infestation. Therefore, the limited search time was run continuously, with searchers requested to move on quickly from each find and continue searching, while the supervisor recorded details of the flagged plant (including any misidentifications). This also reduced the risk of measurement errors arising from frequent stopwatch stopping and starting.

The main limiting resources for scheduling were supervisors, plots and time. While 16 plots were available, many (usually with the same vegetation/yellow flower classification) were close to each other. Searchers might observe and remember the locations of finds made by other searchers in adjacent plots. We therefore scheduled searchers to plots such that they did not search adjacent plots wherever possible. Where adjacent searches occurred, searchers were allocated to the neighbouring plot many hours later. The clustering of plots by type allowed for a schedule with moderate travel distances, a full range of plot types visited throughout the day, and large time lags between replicate searches.

Searchers were available to participate in the experiment over a five day period and there were sufficient searchers available to schedule new participants every half day. However, we elected to involve half as many participants for a full day each. This allowed us to measure the potential effect of fatigue over the day, and increased active search time and data collection by reducing the time spent on travel to the experiment site and inducting new searchers.

We expected that variations in vegetation would influence the speed at which searchers were able to thoroughly search each plot. However, we did not alter the permitted search time, 20 minutes per plot, to accommodate vegetation. We did not want searchers to notice any pattern in variable instructions that might alter their search behaviour.

Trampling of vegetation, particularly in the neighbourhood of targets, was a high risk in this experiment. Scribes were requested to remain outside the plots whenever possible, and searchers were asked to tread lightly. The search schedule permitted that four plots per day could be rested. Nevertheless, some vegetation trampling was unavoidable as we sought sufficient replicates for analysis.

DISCUSSION AND CONCLUSIONS

While conducting a detection experiment is resource intensive, it offers the best opportunity to assess the capability of plant surveys. Such assessment allows us to ensure that sufficient resources are allocated to detect infestations where they occur.

A detection experiment should begin with a clear research question, and identify the variables that are likely to influence weed detection. These may relate to target plant characteristics, the surrounding environment, or the searchers' attributes. Wherever possible, these variables should be controlled, varied and replicated across the experiment. Statistical principles of stratification and randomisation should also be followed.

Nevertheless, the challenge of simulating a realistic search environment and the associated logistical constraints are sometimes at odds with these principles. Issues of realism, feasibility and scientific rigor must be assessed and traded against each other case by case, bearing in mind the purpose of the detection experiment and the survey environment it aims to address.

ACKNOWLEDGMENTS

This project was supported by an Australian Research Council Linkage Project (LP100100441), in partnership with the Victorian Department of Primary Industries, Parks Victoria and the NSW Department of Environment & Heritage. We thank Karen Herbert, Iris Curran, Charlie Pascoe, Kim Millers, Nick Williams, Roger Cousens, and the Hawkweed Project Control Group for their support of this experiment in its design phase.

REFERENCES

Chen, G., Kéry, M., Zhang, J. and Ma, K. (2009). Factors affecting detection probability in plant distribution studies. *Journal of Ecology* 97, 1383-1389.

Emry, D.J., Alexander, H.M. and Tourtellot, M.K. (2011). Modelling the local spread of invasive plants: importance of including spatial distribution and detectability in management plans. *Journal of Applied Ecology* 48, 1391-1400.

Garrard, G.E., Bekessy, S.A., McCarthy, M.A. and Wintle, B.A. (2008). When have we looked hard enough? A novel method for setting up minimum survey effort protocols for flora surveys. *Austral Ecology* 33, 986-998.

Hauser, C.E. and McCarthy, M.A. (2009). Streamlining 'search and destroy': cost-effective surveillance for invasive species management. *Ecology Letters* 12, 683-692.

McCarthy, M.A., Moore, J.L., Morris, W.K., Parris, K.M., Garrard, G.E., Vesk, P.A., Rumpff, L., Giljohann, K.M., Camac, J., Bau, S.S., Friend, T., Harrison, B. and Yue, B. In press. The influence of abundance on detectability. *Oikos*.

Mehta, S.V., Haight, R.G., Homans, F.R., Polasky, S. and Venette, R.C. (2007). Optimal detection and control for invasive species management. *Ecological Economics* 61, 237-245.

Moore, J.L., Hauser, C.E., Bear, J.L., Williams, N.S.G. and McCarthy, M.A. (2011). Estimating detection-effort curves for plants using search experiments. *Ecological Applications* 21, 601-607.

Regan, T.J., McCarthy, M.A., Baxter, P.W.J., Panetta, F.D. and Possingham, H.P. (2006). Optimal eradication: when to stop looking for an invasive plant. *Ecology Letters* 9, 759-766.

Rout, T.M., Salomon, Y. and McCarthy, M.A. (2009). Using sighting records to declare eradication of an invasive species. *Journal of Applied Ecology* 46, 110-117.

Williams, N.S.G., Hahs, A.K. and Morgan, J.W. (2008). A dispersal-constrained habitat suitability model for predicting invasion of alpine vegetation. *Ecological Applications* 18, 347-359.

Williams, N.S.G. and Holland, K.D. (2007). The ecology and invasion history of hawkweeds (*Hieracium* species) in Australia. *Plant Protection Quarterly* 22, 76-80.