M80

QUALITY ASSURANCE/QUALITY CONTROL (QAQC) STUDY:
ESTIMATING DETECTION PROBABILITY BY DOUBLE SAMPLING
PRELIMINARY RESULTS

PREPARED FOR:

Alameda County Community Development Agency
224 W. Winton Avenue, Room 110
Hayward, CA 94544
Contact: Sandra Rivera
510.670.6526

PREPARED BY:

ICF International
630 K Street, Suite 400
Sacramento, CA 95814
Contact: Douglas Leslie
916.737.3000

July 2011

ICF International. 2011. *Quality Assurance/Quality Control (QAQC) Study: Estimating Detection Probability by Double Sampling, Preliminary Results*. July. (ICF 00904.08.) Sacramento, CA. Prepared for Alameda County Community Development Agency, Hayward, CA.

Contents

Tables and Figures

Tables

Figures **Follows Page**

Acronyms and Abbreviations

APWRA	Altamont Pass Wind Resource Area
BLOBs	Base Layer Operating Group Boundaries
CDP	cumulative detection probability
km	kilometers
km^2	square kilometers
m	meters
QAQC	quality assurance / quality control
SRC	Scientific Review Committee

Quality Assurance/Quality Control (QAQC) Study, Estimating Detection Probability by Double Sampling, Preliminary Results

Introduction

The proliferation of wind generation facilities in the United States, and in particular in California, has led to the widespread need to monitor the effects of operating wind turbines on populations of birds and bats. In California, 1–3 years of postconstruction monitoring is typically required by regulatory agencies and land use authorities to determine if actual impacts are in-line with impacts predicted during the environmental review process. This monitoring has most often been accomplished by regularly searching operating turbines within a fixed search area for avian and bat fatalities.

However, the number of fatalities detected during carcass surveys does not equal the actual number of fatalities because some unknown proportion of birds killed by turbines is never observed. Surveyors can miss carcasses that are present within the search area, or carcasses can be removed from the search area during the interval between deposition and the survey, and thus are not available to be detected. To develop a statistically supportable estimate of the actual number of birds killed, the probability of detecting a fatality must be determined.

The customary approach to estimating turbine-related avian fatalities has been to split detection probability into these two components—searcher efficiency and carcass removal—and attempt to estimate these two probabilities separately using searcher efficiency trials and carcass removal trials.

This approach is biased to an uncertain degree. Carcass removal trials are conducted with unknown searcher efficiency, or possibly with searcher efficiency that is different from what would occur during standard searches. The results may differ by the type of evidence monitored (e.g., feather piles, carcasses, parts of animals; see M32), and it is unknown if this is due to variability in searcher efficiency during the carcass removal trial or due to actual differences in removal rates. In addition, there is some evidence that detection probability may change through time in a non-directional way. For example, variable field conditions might enable evidence to be easily detected one day but not the next; similarly, scavengers can move or alter the evidence such that detectability changes, or they can transport evidence in and out of the search area (ICF International 2010).

Searcher efficiency may vary substantially from observer to observer and may have high error from observation to observation (see Collier et al. 2007). Error or variance in one or both components of detection probability will substantially bias the results (White 1998). Compound models of detection probability can be developed to address these issues, but the assumptions of these models are violated if their components are variable or biased (McAllum 2005). We therefore designed and implemented a quality assurance and quality control (QAQC) study—hereinafter referred to as the *QAQC Study*—to determine if using a double-sampling approach to fatality monitoring could provide an estimate of detection probability that is not confounded by these issues.

Double-sampling consists of conducting a survey and conducting QAQC surveys for a subsample of monitored locations (Bart and Earnst 2002). The technique may involve a regular survey followed by an intensive survey. However, the efficiency of the two surveys does not have to be fixed or predetermined to develop a ratio estimator such as detection probability (Royal and Cumberland 1981).

There are a number of possible approaches to developing estimates of detection probability (McCallum 2005). For example, it is theoretically possible to monitor or evaluate the interactions between searcher efficiency and carcass removal rates and all the factors that result in potential bias of the estimates. However, directly monitoring detection probability using a double-sampling approach appears to be a more straightforward approach.

We conducted our study in the Altamont Pass Wind Resource Area (APWRA) as part of a large-scale, ongoing monitoring effort that currently consists of searching approximately 1,000 turbines with a 30–35 day search interval.

Study Area

The APWRA is located in central California approximately 90 kilometers (km) (56 miles) east of San Francisco (Figure 1). There have been as many as 400 wind turbines currently permitted within the APWRA, distributed over 150 square kilometers (km^2) (50,000 acres) of rolling hills and valleys dominated by nonnative annual grassland.

The APWRA supports a broad diversity of breeding, migrating, and wintering bird populations that regularly move through the wind turbine area (Orloff and Flannery 1992). Diurnal raptors (eagles and hawks), in particular, use the prevailing winds and updrafts for soaring and gliding during daily movement, foraging, and migration. Birds passing through the rotor plane of operating wind turbines are often killed. Multiple studies of the avian mortality rates in the APWRA indicate that golden eagles, red-tailed hawks, American kestrels, burrowing owls, barn owls, and a diverse mix of other non-raptor species are killed each year by collisions with turbines (Howell and DiDonato 1991; Orloff and Flannery 1992; Howell 1997; Smallwood and Thelander 2004).

The avian fatality-monitoring program in the APWRA has been conducted continuously since its initiation in 2005 under a sampling regime approved by the Scientific Review Committee (SRC).

Methods

To implement the QAQC Study, the APWRA field crew was broken into teams with different roles and assignments. Search teams were typically deployed in groups of two or three. Team assignments were adjusted each rotation (or occasionally more frequently due to logistical considerations) to minimize systematic bias that could result from fixed team assignments, as well as to account for human resource and logistical requirements. The field supervisor was responsible for managing the double-sampling schedule and protocols.

Teams were *blind* to one another's schedules or results except for rare occasion where logistics required substitution of team assignments or augmentation of teams. The APWRA has been divided into Base Layer Operating Group Boundaries (BLOBs), which are groups of similar turbines in a similar geography for purposes of stratification and allocation of sampling effort. Each team was assigned a set of strings (by BLOB) to conduct the *primary* search (see below for a discussion of

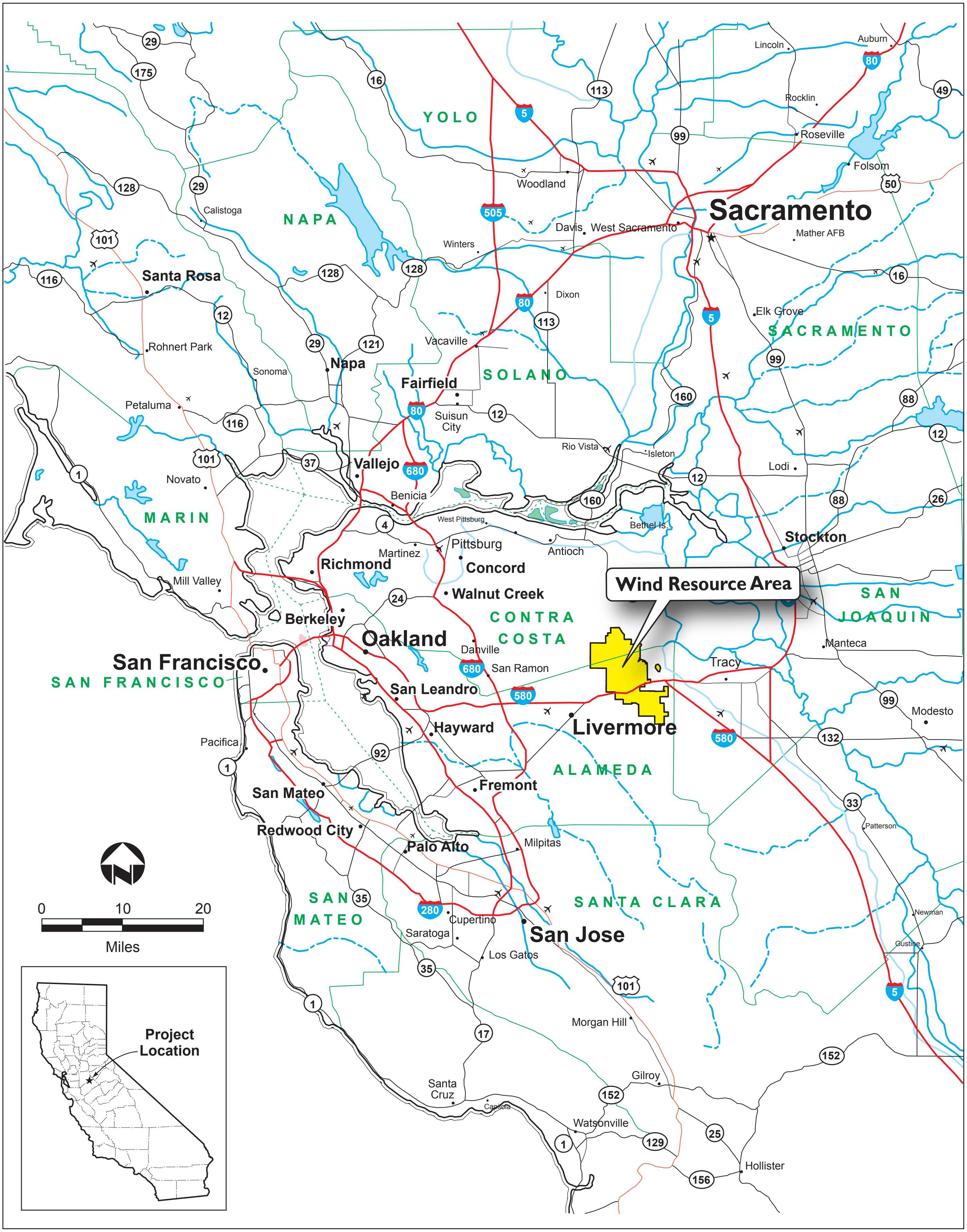


Figure 1
Location of the Altamont Pass Wind Resource Area

Graphics/Projects/00904.08 Altamont Pass/Avaian Fatality Monitoring (11-10) SS

search types) for a rotation (defined as one pass through the complete set of monitored turbines). The searches were scheduled so that each team could move across the APWRA in an efficient manner and complete a primary search of all monitored turbines within an approximately 30–35 day interval.

Several types of searches were conducted as part of the regular monitoring effort, with additional search types implemented as part of the QAQC Study. These search types are defined in Table 1.

Table 1. Types of Searches Conducted in the APWRA QAQC Study

Search Type	Definition	Search Order
Clearing search	A search at turbines that have not been surveyed in more than 90 days, and which might include placed or naturally found bird fatalities left for detection by subsequent searches	0
Incidental discovery	A detection outside of the standard search procedure	0
Wildlife Response and Reporting System	A detection by owner/operators of turbines	0
Pre-search	A search by a supervisor prior to a primary search which may leave placed or naturally found bird fatalities	1
Primary search	A standard search	2
Secondary search	A standard search that follows a primary search within the standard 35-day rotation	3
Post-search	A search by the field supervisor or assistant after a primary or secondary search	4
Fatality check	An examination of a known fatality by the field supervisor or assistant	4

For each rotation a set of turbines was selected for double-sampling during that interval. The number of turbines to be double-sampled was dependent on the workload and available resources. We initially attempted to double-sample approximately 25% of the monitored turbines.

Clearing searches were conducted at any turbines that have not been surveyed in the past 90 days. For the QAQC Study, the field supervisor also conducted clearing searches at all turbines that were scheduled for double surveys during the first search rotation (October 15–November 15, 2010).

The search schedule was randomized so that a variety of intervals ranging from 1 to 30 days between the primary and secondary searches could be implemented during each rotation. However, constraints were placed on the randomization so that a disproportionately high number of secondary searches occurred within 1–2 weeks of the primary search[1]. The secondary search was conducted using the same field protocols as the first, and the second team was blind to the fact that it was conducting a secondary search and to the results of the primary search.

Searchers conducting the clearing, pre-, and primary searches left detected fatality remains in the field to provide secondary searchers with the opportunity to detect those fatalities. Secondary searchers also left fatality remains in the field because they were blind to the fact that they were conducting a secondary and not a primary search. Secondary searches were also exposed to fatalities that were deposited in the interval between the primary and secondary search, and left

---

[1] See the discussion section for details. The current schedule is emphasizing longer intervals to fill this data gap.

those remains in the field as well so they could potentially be collected and then redistributed by the field supervisor (see below).

The schedule was designed to allow the field supervisor to conduct pre-searches at approximately 5% of all turbine searches and approximately 50% of the searches that had secondary or post-searches associated with them (hereafter called *QAQC searches*). The pre-search provides an estimate of the number of fatalities that were available for detection before the primary search. In addition it allowed the field supervisor to actively manage the artificial placement of fatalities at sites with no fatalities. Locations for pre-searches were a randomly selected subset of the double-survey locations for each rotation.

The schedule was also designed to allow the field supervisor to conduct a post-search at approximately 5% of all turbine searches, after 50% of the double-searches, and at all turbines where a fatality was available for detection after the secondary search. Post-searches were conducted approximately 1 day after the last search whenever possible. The locations receiving post-searches by the field supervisor included sites where a fatality was detected on one search but not on a subsequent search, sites where no fatalities were detected, and sites where fatalities were detected by all searches in a rotation.

If a fatality was detected during a pre-search or a primary search but not subsequently detected, the field supervisor conducted a post-search on the subsequent day to determine if the fatality was available for detection. In cases where a fatality was documented during the pre-search but no fatalities were detected during subsequent searches, the field supervisor conducted a post-search to confirm that no fatalities were available for detection and missed. In cases where fatalities were detected by all searches, the field supervisor conducted a post-search to confirm that no additional fatalities were available for detection and missed, and that the condition and biology of the fatality was accurately documented. In cases where the field supervisor could not conduct the post-search due to logistical constraints, the field supervisor or an assistant conducted a fatality check. The fatality check did not include a complete search, but rather a determination of whether or not a known fatality was still present and detectable.

The primary and secondary searches produced detections of fatalities. Results were processed by the field supervisor using a set of data models (Microsoft Excel and web applications) that allowed detections to be assigned to single fatality numbers—to avoid double counting—based on the cumulative results of all searches. The results were used to derive an estimate of detection probability for different fatality types (e.g., feather spots, carcasses), bird groups (i.e., raptors or non-raptors by size), intervals, and searchers.

All fatalities less than 90 days old (i.e., not notably aged) that were detected during pre-searches, primary searches, and/or secondary searches were left in the field to support the double-blind and double-sampling aspects of the program. The field supervisor processed and sometimes removed carcasses found during post-searches. All other remains were processed and removed by search crews under a schedule that was managed by the field supervisor based on travel time and efficiency, and that prioritized fresh fatalities, whole carcasses, or remains that were from species or bird groups that are less common or of higher priority (especially small birds and small raptors). All extant remains were eventually brought back to the field station and archived.

Fatality Placement

The fatality monitoring and QAQC Study design requirements made it impossible to predict a-priori whether a sufficient number of fatalities would naturally be available for detection and/or re-detection at the double sampled sites for a given rotation or season. Therefore, it was necessary to supplement the program with volitionally placed fatalities. The vast majority of these carcasses were fatalities found during regular searches conducted as part of the monitoring program in the APWRA. Fatalities of sufficient quality were collected in the field at the request of the field supervisor, or were collected by the field supervisor when possible. These were held in a freezer, defrosted, and placed onsite at a random set of turbines scheduled to receive double searches.

The field supervisor was responsible for placing fatalities during the pre-searches. Whenever a placement was made, the field supervisor conducted a pre-search to make sure that no additional (and possibly confounding) fatality remains were present. Fatalities were placed within the search area at a random distance and bearing from the turbine. The location and condition of the remains were documented using a modified fatality sheet and a Trimble Yuma GIS data entry system. Placed fatalities were recorded in the APWRA database as a specific fatality type so they would not be used in fatality estimates.

The specific remains placed in the field were selected non-randomly. The field supervisor continually reviewed the types of fatalities detected in the rotation and season, and selected remains that were both available and rare or missing from the double-sampled sites. The goal was to achieve 30 samples per season for each bird group, divided between feather spots and carcass remains.

Analytical Approach

Not all fatalities documented as part of the study were included in the analyses of detection probability (Table 2), although some fatalities excluded from the analysis of detection probability are included in estimating the total number of fatalities APWRA-wide. For example, we did not include fatalities that could not be assigned to a taxonomic group or size class, carcasses that could not be assigned to a turbine or turbine string, or carcasses that were found greater than 125 meters (m) from a turbine. However, we did include fatalities that had a cause of death other than turbine related, incidental finds other than golden eagles, and placed fatalities[2]. We used filtered detections to estimate searcher efficiency, simple and cumulative detection probabilities, and to test adjustment factors based on our estimates of simple and cumulative detection probabilities.

---

[2] These records are excluded when generating the fatality estimates, but included for evaluating the QAQC Study because they meet the criteria of valid detections. Golden eagle remains are removed from the field when detected incidentally and cannot be included in the QAQC Study. In this report we did not analyze QAQC outcomes from incidental detections, but subsequent reports will include this information.

Table 2. Criteria Used to Include or Exclude Records for Evaluating QAQC Results

Status ID	Included in Detection Probability Analyses	Description
0	N/A	Not reviewed
1	Yes	Valid turbine-related fatality
2	No	Fatality cannot be assigned to an animal group or size
3	No	Does not have an associated turbine ID
4	No	Is not within 125 m of the turbine
5	Yes	Cause of death is not turbine-related
6	Yes	Is >90 days dead, aged, or unknown
7	Yes	Incidental or WRRS find except golden eagles
8	No	Injured fatality except golden eagles
9	No	Fatality without a valid search (search ID) or valid detection
10	Yes	Offsite fatality used for placement only
11	Yes	Placed

Searches were assigned to *QAQC search sequences*, which consisted of one or multiple non-primary searches surrounding a primary search (Table 3). These include sequences with pre-searches, post-searches, a secondary search, or some combination of searches. Detections from these searches were used in a variety of combinations to estimate the following parameters.

- *Searcher efficiency*: the probability of a searcher detecting a carcass given that it is present and available for detection.
- *Simple detection probability*: the probability of detecting a fatality across an interval, or the probability a single fatality would remain and be detected.
- *Cumulative detection probability*: the detection probability for an interval, which incorporates all sources of error. Essentially equivalent to the adjustment factor and analogous to the cumulative probability of a carcass remaining multiplied by searcher efficiency for a given bird group or size class.
- *Aggregate detection probability*: either simple or cumulative detection probability or the combination of searcher efficiency and cumulative percent of carcasses remaining from a carcass removal trial.

Detections were also used to test different adjustment factors for various combinations of search sequences. Search sequences were treated as independent data because they were conducted using appropriate randomization routines and blindness.

Table 3. Types of Search Sequences Used in the Analyses of Detection Probability

Search Sequence Type	Search Sequence Type Description
Primary only[a]	Includes only a primary search
Not a valid sequence[a]	Missed a primary search or primary search interval
Primary and secondary	Includes a primary and secondary search
Primary and post-	Includes a primary and post-search
Pre-, primary, and post-	Includes a pre- (or clearing), primary, and a post-search
Primary, secondary, and post-	Includes a primary search, secondary, and post-search
Pre-, primary, and secondary	Includes a pre- (or clearing) search, primary, and secondary search
Pre-, primary, secondary, and post-	Includes a pre- (or clearing) search, primary, secondary, and post-search
Primary and multiple secondaries	Includes a primary search and multiple secondary searches

[a] Not used in the QAQC analysis.

Estimating Searcher Efficiency

In estimating searcher efficiency, the double-blind design and supervisor surveys provide several opportunities to determine if searchers found bird remains that were known to be present. The technique is similar to double-blind sampling used in avian population monitoring (Collins 2007) where a first sample is followed by a second (usually more intensive) sample to estimate the ratio of misses and bias. In the case of fatality monitoring, all primary and secondary searches were intensive and complete and provide a good estimator of searcher efficiency.

A fatality placed or found during the pre-search is available for detection by a primary search if it is discovered on the primary search (because it was detected) or discovered later by the secondary search or a post-search. Similarly, a fatality detected on either a pre-search or primary search, and subsequently detected on the secondary search, post-search, or next primary search is available for detection on the subsequent secondary search[3]. Detections from pre-and post-searches were used only to confirm the presence of a fatality, but not to estimate searcher efficiency, as those searches are not blind and do not follow the same protocols as primary or secondary searches.

Searcher efficiency was calculated as the simple percentage of fatalities detected and known to be available for detection. Fatalities detected on the pre- and post-search were available for detection during both the primary and secondary searches, and provide an estimate of the variability in searcher efficiency. Variability in searcher efficiency was calculated as the percentage of fatalities detected by both primary and secondary searches that were available for detection by both search events.

---

[3] Incidental and WRRS detections left in the field are believed to be available for detection under similar criteria, and will be included in future versions of this report.

Estimating Detection Probability

To estimate aggregate detection probabilities we evaluated the probability of re-detecting a fatality over time. Initial detections (and placements) on a pre-search or clearing search were evaluated to determine if they were detected by the primary search. Initial detections (and placements) made on the pre-search, clearing search, or primary search (or any appropriate incidental detections) were evaluated to determine if they were detected by the secondary search[4]. Subsequent primary searches, while useful in estimating the persistence of a fatality in the environment, were not blind and therefore were not used to estimate detection probability.

We used univariate survival analysis based on initial detections, subsequent searches, and subsequent detections in a search sequence to estimate Kaplan-Meier survival parameters (Kaplan and Meier 1958). Time to failure was estimated based on the time between initial detections and subsequent searches in a sequence. Detection at the end of the sequence was used to censor fatalities, allowing them to “survive” in the model for the interval between detections. We used lognormal fit to estimate parameters of a survival curve that estimates detection probability directly, incorporating both searcher efficiency and the probability that a fatality was removed from the search area between initial detections and subsequent searches (Kalbfleisch and Prentice 1980).

We developed aggregate cumulative detection probability (CDP) curves by averaging the probability of detection over all days in the interval *(I)* of interest using the equation

$$R_{ag} = \frac{\sum_{I}^{i=1} R_{agi}}{I \text{ x } 100}$$

where R_{ag} is the aggregate probability of detecting a fatality for the ith day between detection and the subsequent search.

Adjusted QAQC Fatality Estimates

We tested the ability of various CDP models to estimate known numbers of fatalities from subsequent blind secondary searches by dividing the number of secondary detections by R_{ag} for the average interval. To test the accuracy of the fitted CDP model, we adjusted fatalities for the same populations used to develop the model. For example, we estimated the number of placed fatalities (a known population) from the number of placed fatalities subsequently detected. In addition we tested each CDP model against populations that did not contribute substantially to its development. For example, we estimated the number of all primary detections by adjusting the number of secondary detections using CDP models developed only from fatalities 0–3 days of age (termed *fresh*), less than 8 days of age, less than 30 days of age, and placed fatalities each in succession. For comparison we also adjusted all known populations of fatalities using adjustment factors derived from Smallwood (2007) to determine which models provided the most accurate estimate.

---

[4] The primary search and secondary search represent independent samples of each fatality and can be used as independent data to increase the sample size. Subsequent versions of this report will include all combinations of primary and secondary detections.

Results

A total of 252 primary searches—or approximately 18%—had QAQC searches associated with them (Table 4); of these, 108 (43%) contained a pre-search and 144 (54%) contained a post-search. An additional 21 sequences were not included in the QAQC analysis because they consisted of a pre- or post-search only or a secondary search only. These were typically conducted at strings that were removed from the design in the middle of the bird year.

Table 4. Sample Size, Number of Events, and Intervals between Events

Search Sequence Types	Number of Sequences	Average Number of Events in Sequence	Average Interval Between Events	Average of Total Sequence Interval	Average of Interval to Next Primary
Not used for QAQC					
Not a valid sequence	21	1.3	1.0	2.4	2.4
Primary only	1,094	1.0	0.0	0.0	26.8
With pre-searches					
Pre-, primary, and post-	2	3.0	3.8	11.5	46.5
Pre-, primary, and secondary	10	3.0	8.9	26.6	35.9
Pre-, primary, secondary, and post-	63	4.0	3.4	13.6	41.0
Pre- and primary	33	2.0	7.0	13.9	15.2
Total with pre-searches	108				
Without pre-searches					
Primary and multiple secondaries	9	6.4	1.4	8.4	30.8
Primary and post-	11	2.0	3.1	6.2	26.1
Primary and secondary	79	2.0	4.0	8.0	38.6
Primary, secondary, and post-	45	3.0	4.2	12.5	35.1
Total without pre-searches	144				
Grand Total	1,367	1.4	0.8	2.2	27.9

Detections were distributed broadly across bird size (Table 5) and species (Table 6). Detections included 53 placed birds and 813 subsequent valid detections in the field. The average estimated age at time of detection for all birds was 34.5 days, with considerable variability across species, bird size, and volitionally placed versus non-placed birds (Table 6). Slightly less than half of all primary detections had associated QAQC searches (Table 7). More than two-thirds of all post-searches consisted of full searches of the string versus fatality checks.

Table 5. Fatality Detections and Placed Carcasses Included in the Analyses of Detection Probability by Size Class

	Size			
Detection Status	Large	Medium	Small	Grand Total
Valid detection	325 (40%)	207 (25%)	281 (35%)	813 (94%)
Placed bird	26 (49%)	5 (9%)	22 (42%)	53 (6%)
Grand total	351 (41%)	212 (24%)	303 (35%)	866

Table 6. Fatality Detections Included in the Analyses by Size Class and Species

Size	Common Name	Number of Detections	Average of Estimated Days Dead
Found Birds			
Large	American crow	5	72.0
Large	barn owl	64	32.3
Large	California gull	16	21.0
Large	common raven	22	27.6
Large	ferruginous hawk	6	8.3
Large	golden eagle	19	58.7
Large	great-horned owl	13	31.2
Large	red-tailed hawk	99	28.8
Large	turkey vulture	3	36.3
Large	unidentified buteo	11	78.8
Large	unidentified gull	40	28.7
Large	unidentified large bird	27	76.0
Large Total		325	35.9
Medium	mallard	9	39.5
Medium	medium bird	26	69.3
Medium	mourning dove	6	19.0
Medium	northern flicker	5	45.0
Medium	rock pigeon	147	29.1
Medium	unidentified dove	10	48.2
Medium	unidentified duck	3	60.0
Medium	white-tailed kite	1	45.0
Medium Total		207	38.6
Small	American coot	4	10.5
Small	American kestrel	41	26.6
Small	Brewers blackbird	7	26.8
Small	burrowing owl	31	41.2
Small	European starling	65	35.9
Small	horned lark	3	2.0
Small	killdeer	1	
Small	loggerhead shrike	5	
Small	small bird	41	51.3
Small	unidentified blackbird	2	
Small	unidentified bluebird	5	90.0
Small	western meadowlark	76	25.7
Small Total		281	34.4

Size	Common Name	Number of Detections	Average of Estimated Days Dead
Placed Birds			
Large	barn owl	5	18.2
Large	California gull	2	10.5
Large	common raven	3	17.7
Large	ferruginous hawk	1	2.0
Large	great-horned owl	1	2.0
Large	red-tailed hawk	12	5.6
Large	turkey vulture	1	45.0
Large	unidentified gull	1	6.0
Large Total		26	12.0
Medium	mallard	1	
Medium	mourning dove	1	
Medium	rock pigeon	3	3.3
Medium Total		5	3.3
Small	American kestrel	7	14.8
Small	Brewers blackbird	3	
Small	burrowing owl	3	19.0
Small	European starling	2	10.5
Small	savannah sparrow	1	2.0
Small	western meadowlark	6	10.5
Small Total		22	13.1
Grand Total		866	34.5

Table 7. Number of Detections by Search Type and Size Class for Sequences Used in the QAQC Analysis

Search Type	Search Sub-Type	Small	Medium	Large
Pre-	Clearing	13	10	29
	Pre-search	25	5	28
Primary	QAQC primary	45	33	53
	Primary only	120	103	112
Secondary		33	22	51
Post-	Post-search	37	19	29
	Fatality check	11	2	10

Searcher Efficiency

Small Birds[5]

The percentage of available small birds found by secondary searches was lower than those for primary searches by 9%, but this difference was not statistically significant based on a simple ANOVA (R^2=0.01, p=0.54). The percentage of available placed birds detected was statistically similar for both primary searches and secondary searches based on a simple ANOVA (R^2=0.01, p=0.83, Tables 8 and 9). However, despite the fact that searcher efficiency was similar between primary and secondary searches, only 14% percent of small birds known to be present were detected by both the primary and secondary searches.

Table 8. Number of Detections Placed and Found by Bird Group and Search Type for Placed Birds

	Large Non-raptors	Large Raptors	Medium Raptors	Medium Non-raptors	Small Non-Raptors	Native Small Raptors	Nonnative Medium Non-Raptors	Nonnative Small Non-Raptors
Placed during pre-search	6	20	0	2	11	10	3	2
Primary search	3	13	0	2	4	10	3	0
Secondary search	2	20	0	1	2	5	4	0
Post-search	2	13	0	1	8	3	2	0

Table 9. Searcher Efficiency Estimates by Bird Size Based on Fatalities Available for Detection

		Small					Medium[a]			Large[a]		
Found by		n	Percent	Std Dev	R^2 with interval	P	n	Percent	Std Dev	n	Percent	Std Dev
Primary	Placed	15	.40	.51	0.04	0.53						
	Not placed	7	.57	.53	0.35	0.16						
	Both	22	.50	.51	0.10	0.17						
Secondary	Placed	13	.39	.51	0.15	0.19						
	Not placed	14	.40	.51	0.01	0.69						
	Both	27	.41	.50	0.02	0.46						
Primary or secondary		47	.57	.51								
Primary and secondary		14	.14	.36								

[a] Results will be included in the next draft.

---

[5] Results for other size classes will be included in the next draft.

The probability of detecting an available small bird was not significantly correlated with the length of the preceding interval between initial detection and subsequent survey for either primary, secondary, or combined searches based on simple linear regression (Table 9). A nominal logistic regression between search interval length and searcher efficiency was significant for placed small non-raptors, and marginally significant for non-placed small raptors (Figure 2).

Detection Probability Small bird simple detection probabilities varied as a function of search interval and across bird groups. Aggregate detection probabilities were higher for non-placed birds than for volitionally placed birds. Small raptor simple ratio detection probabilities were higher than non-raptors for placed birds, but similar for non-placed birds (Figure 3).

Simple aggregate detection probabilities were lowest for fresh birds, and increased with the age of the sample used to generate the curve (Figure 3). The lognormal fit was significant and robust for all curves (Table 10). Data used to generate curves for fresh and newer fatalities was biased towards shorter intervals, and the curves do not yet appear to be robust out to 30 days since they cross the X-axis, suggesting a negative detection probability[6]. The same trends appeared to be true for cumulative detection probabilities (Figure 4).Cumulative detection probabilities were lowest when calculated using fresh birds, and increased with the age of the records in the sample used to generate each model (Figure 4 and Table 10). All curves suggest a higher CDP for days 1–2 than the CDP's calculated from Smallwood (2007). However, the current CDP models for fresh fatalities and fatalities less than 8 days old were derived from short intervals and suggest a lower detection probability at the 30-day mark (Figure 4).

Table 10. Curve Parameters for the Aggregate Simple and Cumulative Detection Probabilities for Each Small Raptor Sample Population

Aggregate Simple			Cumulative		
Sample	Lognormal Fit	R^2	Sample	Lognormal Fit	R^2
Fresh (0–3 days)	y = -0.243ln(x) + 0.7814	0.93	Fresh (0–3 days)	y = -0.193ln(x) + 0.8429	0.99
Less than 8 days	y = -0.262ln(x) + 0.8522	0.95	Less than 8 days	y = -0.208ln(x) + 0.9185	0.99
Less than 30 days	y = -0.282ln(x) + 1.0249	0.98	Less than 30 days	y = -0.224ln(x) + 1.0962	0.99
Placed	y = -0.287ln(x) + 0.9853	0.97	Placed	y = -0.228ln(x) + 1.0579	0.99
All birds	y = -0.285ln(x) + 1.0308	0.98	All birds	y = -0.226ln(x) + 1.1029	0.99
			Smallwood	y = -0.165ln(x) + 0.8177	0.98

Fatality Estimates

Populations of "known" carcasses were assembled for placed carcasses only, fresh carcasses only, carcasses less than 8 days old, carcasses less than 30 days old, all carcasses detected on secondary searches, and all carcasses from double-samples (including primary detections from pre-searches) based on the number of initial (pre- and clearing) and primary detections. We calculated the number of adjusted fatalities for each known population based on the number of secondary detections and the estimated detection probability for the corresponding search interval for each known population. However, when these populations were subdivided by taxonomic and bird size groups, there were insufficient small raptor records to test the detection probability models against known

---

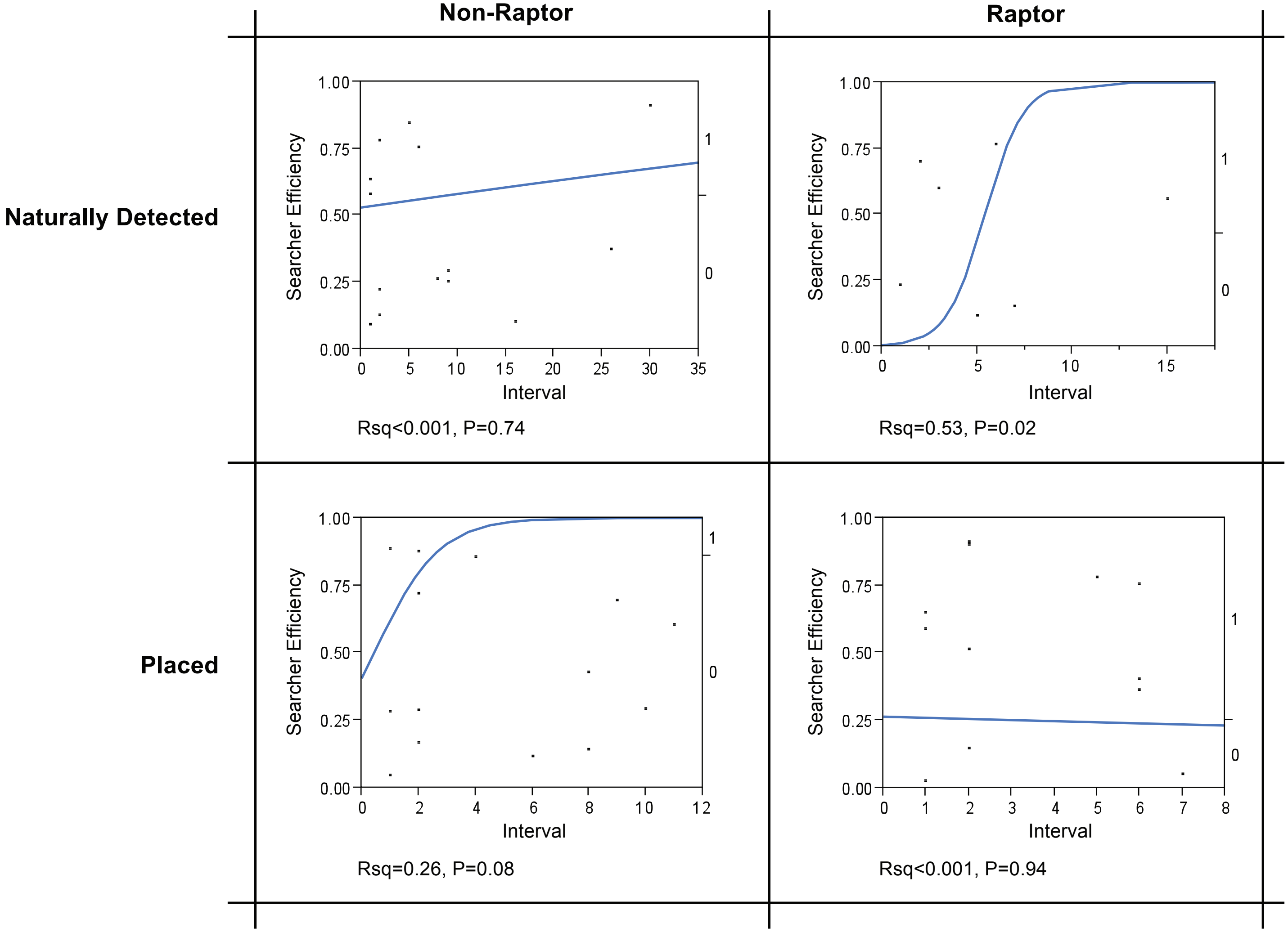
[6] Longer interval data will be available in subsequent versions of this report, and the results are expected to be influenced by those data.

populations of found fresh fatalities or found fatalities less than 8 days old. The relatively short search intervals and small sample sizes made use of cumulative detection probabilities inappropriate; we therefore used simple aggregate detection probabilities to adjust fatalities.

Simple aggregate detection probabilities developed from fatalities less than 30 days old, and from all fatalities, accurately predicted the number of placed small raptors in the “known” population from the number of detections. Fresh fatalities overestimated the population, whereas the adjustment factors derived from Smallwood (2007) underestimated the population (Table 11). Similar patterns were seen when estimating the number of fatalities from the population of carcasses less than 30 days old, all fatalities with secondary searches, and all fatalities from double-sample searches. In each case the adjustment factors derived from Smallwood (2007) underestimated the population. At least one or more of the QAQC-derived models more accurately estimated the actual number of fatalities from subsequent detections. All double-sampled fatalities represented the largest population with the longest interval, and this population was overestimated by all QAQC-derived simple detection probabilities. The fresh fatality CDP most accurately predicted the number of fatalities in this population, underestimating by just 1.6 fatalities.

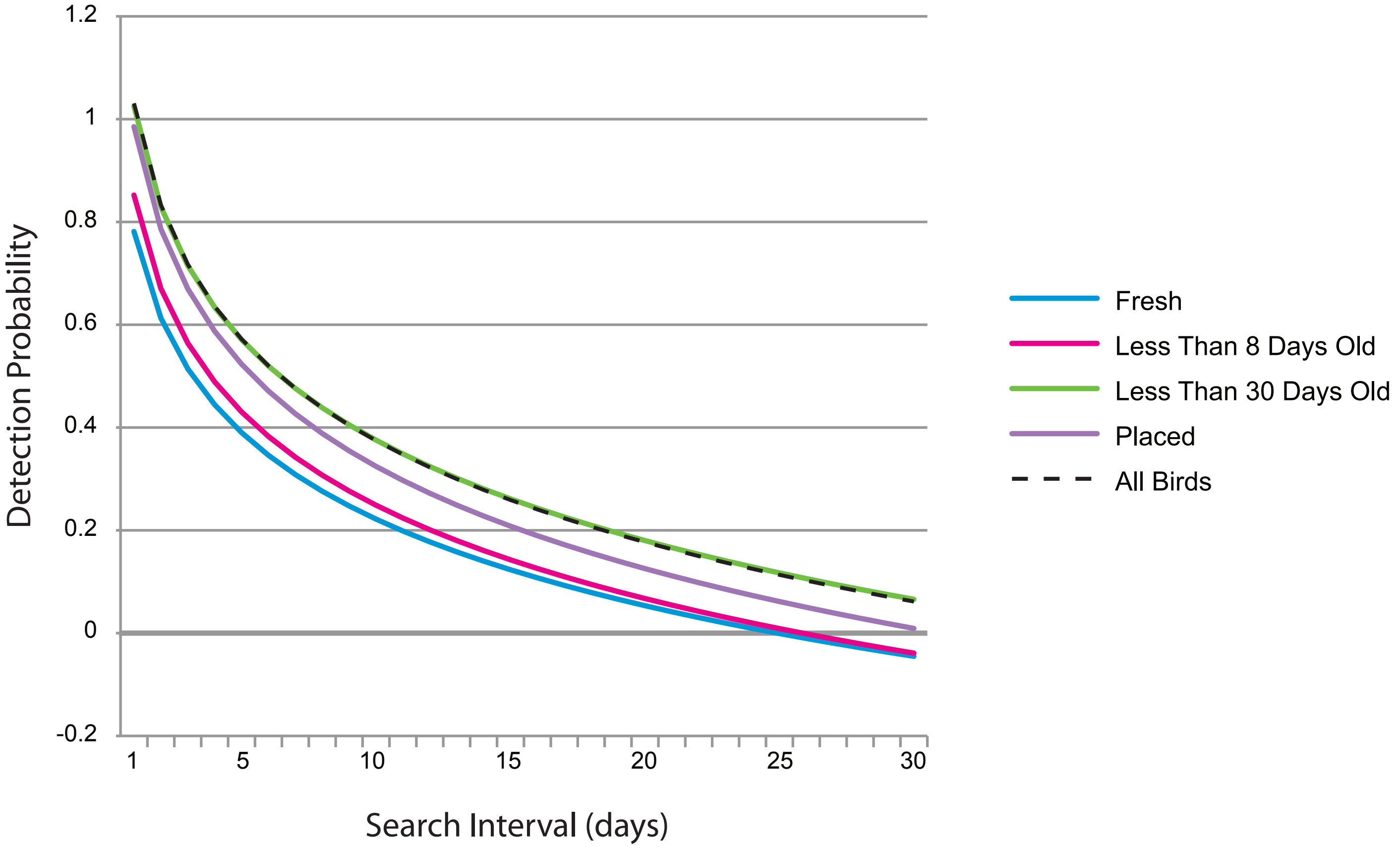
Table 11. Adjusted Fatality Estimates across Various Experimental Intervals Using Multiple Models and the Number of Initial (I), Primary (P), and Secondary (S) Detections and Their Corresponding Search Intervals

		Searches In Sequences			First Detections				Initial to Secondary Interval in Days		Adjusted Estimates from Secondary Detections					
Initial Detection Type	Bird Group	I	P	S	I	P	I+P	S	Mean	Std Dev	Smallwood	Placed	Fresh	< 7	< 30	All
Placed	Raptors	9	9	9	9		9	4	8.2	9.6	8.0	10.3	14.5	13.0	9.1	9.1
	Non-Raptors	8	8	8	8		8	1	6.8	3.4						
Fresh	Raptors	2	2	2	1	1	2	0	2	1.4						
	Non-Raptors	3	5	5	3	2	5	2	6.4	5.0						
Less than 8 days	Raptors	3	3	3	1	2	3	0	3.7	3.1						
	Non-Raptors	3	6	6	3	3	6	2	9.5	8.8						
Less than 30 days	Raptors	10	10	10	8	2	10	3	8.1	9.0	6.0	7.7	10.9	9.8	6.8	6.8
	Non-Raptors	8	11	11	7	4	11	4	7.6	7.0						
All records searched by the secondary	Raptors	14	16	16	11	5	16	6	7.3	7.6	11.5	14.1	19.4	17.5	12.6	12.6
	Non-Raptors	17	27	27	15	12	28	6	11.8	9.9						
All double-samples	Raptors		27	27			27	8	16.8	12.4	21.5	42.2	74.3	63.7	32.9	33.2
											CDP	18.4	25.4	22.9	16.5	16.5
	Non-Raptors		89	89			89	9	24.1	13.5						



Graphics ... 00904.08 (7-11)

Figure 2
Nominal Logistic Regression between Search Interval Length and Searcher Efficiency for Detected and Placed Small Birds



ICF INTERNATIONAL

Figure 3
Small Raptor Simple Aggregate Detection Probabilities

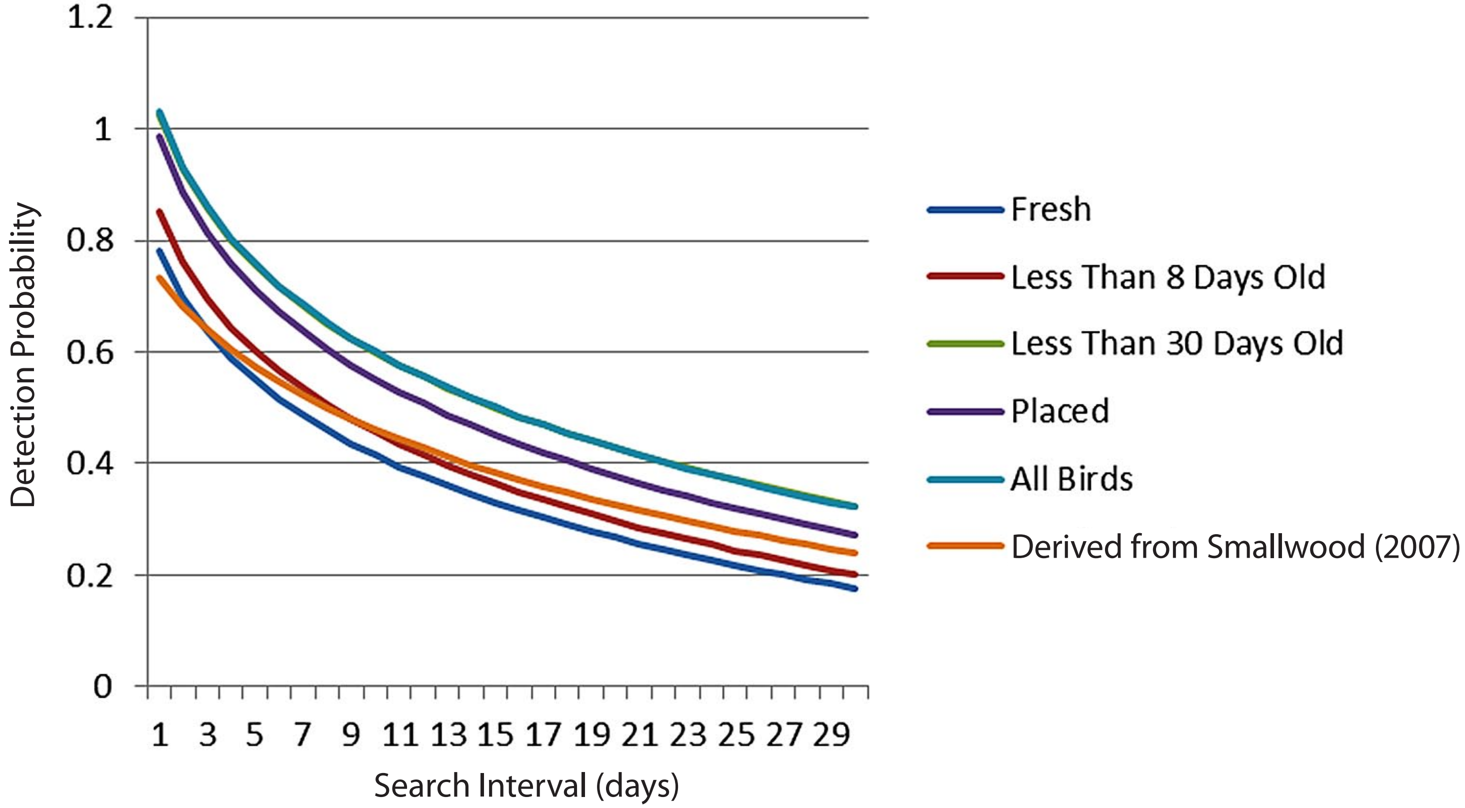


Figure 4
Small Raptor Cumulative Detection Probabilities by Sample

Discussion and Conclusions

The QAQC design was implementable without interfering with the primary search interval. Logistics and person-power limitations resulted in a relatively small number of the more complex search sequences. Simple double-sampling (primary to secondary to next primary) provided a large amount of information and was able to be implemented more easily within the constraints of the ongoing monitoring program. The number of QAQC events and sequences implemented was similar to plan, but the timing of events was biased towards shorter sequence intervals.

Searcher Efficiency

Event-specific searcher efficiency is relatively predictable and was similar to previously published estimates of 51% for small birds (Smallwood 2007). However, variability in small bird searcher efficiency was alarmingly high; primary and secondary searches produced similar results only 14% of the time. Turbine-related fatalities are relatively rare events, and the opportunity to detect a fatality presents itself in only a small percentage of searches. The possibility that searchers may produce highly different results for rare detection opportunities calls into question the reliability of all small bird fatality estimates derived from human visual transects. However, over a large number of events the variance in searcher efficiency may not influence the estimates of detection probability or the estimates of fatality rates.

In evaluating the relationship between searcher efficiency and the length of the search interval we are testing the hypothesis that searcher efficiency decreases over the time that any fatality is in the environment and technically "available and detectable." The null hypothesis is that searcher efficiency is similar for fatalities of any age within the search interval, and that estimates of searcher efficiency are not biased by age. The null hypothesis was supported for all bird groups (with the possible exception of small raptors) based on a simple regression between search interval length and the probability of detecting a known fatality.

However, for small raptors (placed and naturally found) we detected a moderate and perhaps biologically significant relationship, albeit using a small sample size (see Table 4). A logistic fit of search interval length on the categorical determination of detection suggests a biologically significant relationship between search interval and searcher efficiency (R^2=0.13, p=0.08).We believe this relationship should be explored further once the full bird-year data is available.

Cumulative Detection Probability

Preliminary estimates of small raptor CDP were lower than previously used estimates derived from Smallwood (2007) for the short (0–16 day) average intervals we evaluated. CDP estimates based on only QAQC sequences differed—and were typically lower than—estimates that included detection rates for subsequent primary searches. CDP estimates varied based on the type of evidence that was included and, as expected, tended to be higher with increasing age of the evidence.

The sample size for small raptor carcasses estimated to be less than 8 days of age was relatively small. Estimates of CDP derived from fresh fatalities and fatalities less than 8 days of age were higher than previously used estimates derived from Smallwood (2007) in the early portion of the interval, but were lower than those derived from Smallwood (2007) for the latter part of the interval. This suggests that previous estimates of removal rates and/or searcher efficiencies for fresher fatalities were lower than previous estimates for small raptors. The actual CDP-interval relationship may have a more complex shape than previously predicted, which suggest a better approach may be to use a variety of evidence to assemble a final CDP curve.

Adjusted Fatality Estimates

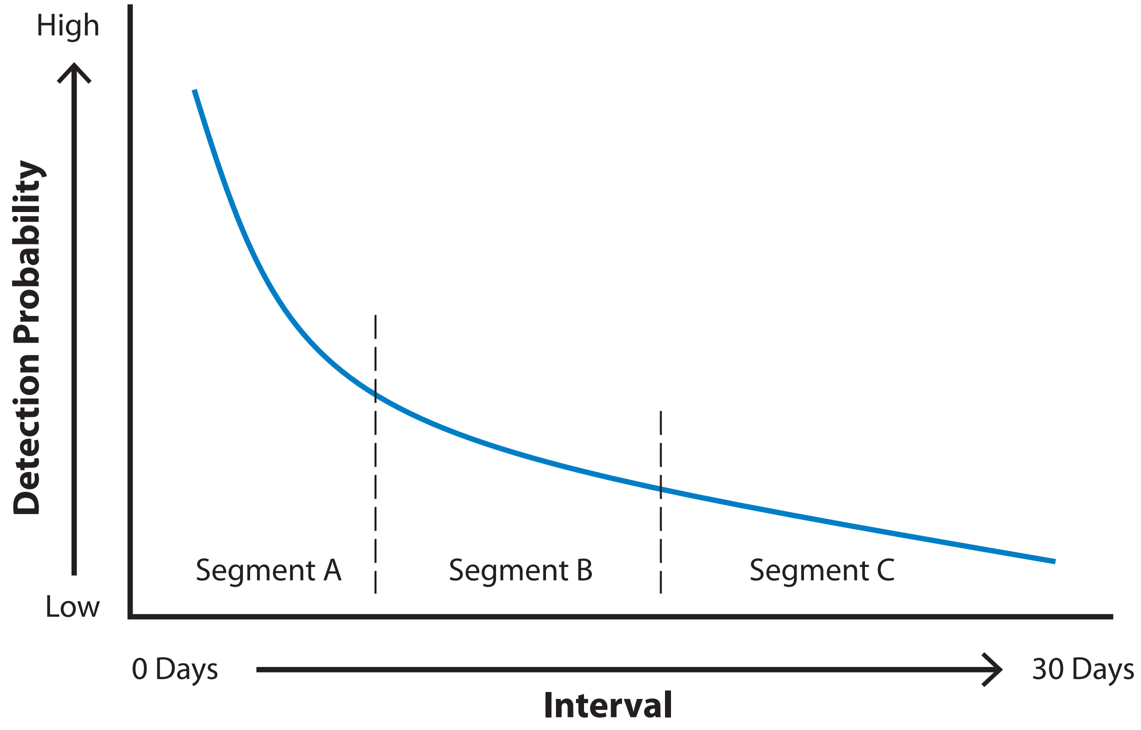
The ability of QAQC-derived CDP estimates to accurately adjust fatalities for known populations of sub-sampled small raptors was specific to the estimates used and the population being adjusted. Within primary interval CDP estimates were more accurate adjustment factors for most of the short-intervals we evaluated. The simple detection probability was a more accurate small raptor adjustment factor than the CDP in many cases, presumably because of the short intervals being evaluated.

Placed birds included mixed-age fatalities, and were best predicted by the simple detection probability for birds less than 30 days old or the detection probability derived from all birds. Fatalities less than 30 days old were double sampled within approximately 1 week, and were most accurately predicted by the simple detection probability for fatalities less than 8 days old. These fatalities, the primary target of a 30-day search interval, were overestimated by the simple detection probability from fresh fatalities, and underestimated by all CDP estimates.

The double-sampled population was most accurately estimated by the simple detection probability for small raptors less than 8 days old, and the simple detection probability for placed fatalities. The all-QAQC population was under-adjusted by estimates derived from Smallwood (2007), the fresh fatality CDP derived from the cross-sequence detections, and the fresh fatality CDP derived from the QAQC sequences. However, the fresh fatality CDP based on QAQC sequences was the most accurate adjustment factor, followed by the simple detection probability for fatalities less than 30 days old.

Based on these preliminary findings we conclude that the relationship between search interval and the cumulative detection probability is more complex than previously described. We present the hypothesis that an approximately 30-day CDP curve consists of three regions that can be accurately predicted to varying degrees based on the QAQC information available (Segments A, B, and C in Figure 5).

Segment A receives fresh small raptor fatalities which are unlikely to persist through a 30-day interval with a frequency of perhaps 3–6% (Smallwood 2006; ICF Jones & Stokes 2009). Fresh fatalities deposited in this “early” segment of the interval are subject to scavenging and environmental factors that influence the removal rate. If a site is searched within 1 week of this interval, the detection probability would be most sensitive to this removal probability, provided that searcher efficiency was sufficiently high and relatively consistent. We hypothesize that the strength of the detection probability–interval relationship is high for a 1-week interval, as evident by the accuracy of the detection probabilities derived from placed fatalities and fatalities less than 8 days old in adjusting the double-sampled population across an approximately 1-week interval. However, all evidence that persists through this segment of the interval will be subject to scavenging and environmental factors and may become increasingly difficult to detect by searchers in the next segment of the interval. Segment A also receives evidence from the previous interval whose



Graphics ... 00904.08(7-11)

Figure 5
Relationship between Cumulative Detection Probability and Search Interval
Depicting Three Hypothetical Regions of the Curve

detection probability is almost entirely determined by searcher efficiency related to evidence that is 0–30 days old[7].

Segment B receives fresh fatalities with an equal likelihood as Segments A and C. In addition it receives fatalities 1–10 days old from Segment A, plus fatalities greater than 10 days old from the previous interval. The likelihood of fresh small raptor evidence persisting from Segment A to Segment B is perhaps 50% (based on Smallwood 2007) or 60% (based on this study[8]).

The probability of detecting older evidence persisting from Segment A to Segment B is almost entirely a function of searcher efficiency. Older evidence in Segment A exists in Segment A because it was missed in the previous interval to some varying and episodic degree. Evidence which persists from the previous interval through to Segment B is likely 14–30 days old, cannot be discriminated from evidence deposited in Segment A, and is less subject to removal. This problem is further confounded during Segment C.

Segment C is the most recent portion of the interval and (across the year) will contain fatalities of varying ages, including fatalities missed in previous intervals. Segment C will contain some portion of 0–10 day old fatalities that are most likely to be detected if they are deposited within 0–3 days of the search. In theory these fatalities are representative of all fatalities deposited in the interval, and it should be possible to adjust the actual population of fatalities based only on this evidence. Because all fatalities are instantaneously fresh when they are deposited, the aggregate probability of detecting a fresh fatality is theoretically a strong predictor of the total number of fatalities deposited throughout the interval.

Segment C also contains fatalities deposited during Segment B, Segment A, and the previous interval. Non-fresh fatalities in Segment C are increasingly sensitive to variability in searcher efficiency depending on their age. As fatalities age they are less likely to be removed, and may be more likely to be missed by the searcher to some, perhaps highly variable, degree. If missed again the oldest evidence from the previous interval may be aging to such a degree that it would be either removed by environmental variables or classified as "aged"—both of which would remove it from the estimate and eliminate its bias in subsequent intervals under the current analysis framework. We suspect that the frequency by which evidence persists across more than one interval and biases the fatality estimate is relatively small.

However, evidence from Segment A and B is now moderately aged, and will be missed by searchers to a highly variable degree. In addition evidence from Segment A and B can be confused with

---

[7] The relationship between searcher efficiency and "removal rates" has been discussed at times in reference to the concept of "double counting." The concern which has been expressed is that the adjustment factor aims to account for fatalities missed in the previous interval, but the actual unadjusted fatality counts might include fatalities which were previously missed.

Bias from cross-interval detections cannot be avoided unless fatality estimates are based only on fresh fatalities, because the age of anything that is not fresh cannot be reliably associated with the current versus the previous interval. Hence it would neither be right nor wrong, but rather a violation of first principles to treat the searcher efficiency probability as if it were disjoint from the "percent remaining" variable, because searcher efficiency influences the "percent remaining" from the previous interval to some unknown degree. It may be possible to develop an estimate of the joint probability of cross-interval searcher efficiency bias on percent remaining from the QAQC data. However we believe that this approach would produce an unreliable adjustment factor due to the rarity of fatality detections and the variability of searcher efficiency documented in this study.

[8] The 0–14 day aggregate detection probability averaged approximately 34% for estimates derived from fatalities that were fresh or fatalities estimated to be 0–8 days of age, divided by our overall searcher efficiency estimate of 57% equals approximately 60%.

evidence deposited fresh during Segment C, except for the most obviously fresh fatalities. This may explain why the simple detection probability derived from fresh fatalities only significantly overestimates the all-QAQC population and the fresh fatality CDP underestimates the population.

The evidence in the all-QAQC approximately 17-day interval comes from a scattered and uneven number of days. The probability of detecting each of these fatalities during a secondary search is specific to the age of the fatality, and thus may not be well suited to extrapolating over long intervals. In a typical monitoring program, this interval would also be influenced by searcher efficiency and the related input of evidence from the previous interval. Variability of searcher efficiency on this evidence is specific to the age and condition of each fatality, and is likely difficult to predict.

It is appealing to consider a detection probability curve based only on fresh fatalities and the detection of fresh fatalities. Previous attempts to develop this approach have been confounded by the variability and rarity by which fresh fatalities are detected. Reducing the search interval to 2 weeks more than doubles the probability of detecting extant fresh fatalities. This is because the probability of removal is lower for the actual day of the search, and perhaps for 1–3 days prior, than for the earlier portions of the interval. The searcher has a high probability of finding a small bird for perhaps three-fourteenths of the interval in a 2-week design, versus three-thirtieths in a 30-day design. Obviously a 1-week design offers further benefits. We believe the power of a 2-week design or a 2-week double-sampling regime should be explored statistically in terms of the relationship between decreasing the number of monitored strings and increasing the precision of the overall program.

On a 30-day interval we hypothesize that the fresh fatality CDP may overestimate the number of fatalities unless variability in searcher efficiency can be addressed. The fresh fatality CDP is highly influenced by the removal rate, disproportionately adjusts older detections which were missed in previous intervals, and does not address variance in searcher efficiency throughout the interval or on fatalities of different age categories. Similarly, we hypothesize that traditional searcher efficiency estimates under estimate searcher efficiency and ignore variability and lack of precision in searcher bias. If a 30–40 day search interval is to be implemented we would recommend the development of an aggregate 30–40 day CDP using multiple models developed from multiple QAQC techniques currently being deployed.

Recommendations

1. Complete the current bird year QAQC program as planned, and leave all fatalities in the field in all cases to allow for subsequent detection by primary searches.
2. Focus QAQC efforts on longer intervals for the remainder of the year to help in the evaluation of the later portion of the CDP-interval relationship.
3. Incorporate incidental and WRRS fatalities into the searcher efficiency and CDP estimates.
4. Continue with placed fatalities, and focus placements towards the beginning of longer sequences.
5. Evaluate the statistical power of the current level of the QAQC design, and conduct a power analysis to determine the level of effort needed to produce CDP estimates.

6. Evaluate the statistical gains and losses of decreasing the overall sample size and decreasing the search interval to 2 weeks for long-term monitoring.
7. Evaluate the financial costs and statistical benefits of incorporating dog searches into the design to minimize searcher bias and variability in searcher efficiency.

References

Bart, J., and S. Earnst. 2002. Double Sampling to Estimate Density and Population Trends in Birds. *Auk* 119(1):36–45.

Collier, B. A., S. S. Ditchkoff, J. B. Raglin, and J. M. Smith, 2007. Detection Probability and Sources of Variation in White-Tailed Deer Spotlight Surveys. *Journal of Wildlife Management* 71(1):277–281.

Collins, B.T. 2007. Guidelines for Using Double Sampling in Avian Population Monitoring. *Auk*. 124(4):1373–1387.

Howell, J. A. 1997. Avian Mortality at Rotor Swept Area Equivalents, Altamont Pass and Montezuma Hills, California. *Transactions of the Western Section of the Wildlife Society* 33:24–29.

Howell, J. A., and J. E. DiDonato. 1991. *Assessment of Avian Use and Mortality Related to Wind Turbine Operations, Altamont Pass, Alameda and Contra Costa Counties, California, September 1998 through August 1989*. Final Report submitted to U.S. Windpower, Inc., Livermore, CA.

ICF International. 2010. Examples of datasheets and photographs associated with the 48-Hour Survey study presented in M32. Available: <ecosystems.icfi.com/windenergy/altamontpass/Photos/Photos.aspx>. Accessed: July 20, 2010.

ICF Jones & Stokes. 2009. *Altamont Pass Wind Resource Area 48-Hour Search Interval Bird Fatality Study (Kestrel and Burrowing Owl Report)*. June. (ICF J&S 00904.08.) Prepared for the Altamont County Community Development Agency, Hayward, CA.

Kalbfleisch, J. D., and R. L. Prentice. *Statistical Analysis of Failure Time Data*. New York: Wiley 1980.

Kaplan, E. L.; Meier, P. 1958. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* 53:457–481.

McCallum, D. A. 2005. *A Conceptual Guide to Detection Probability for Point Counts and Other Count-based Survey Methods*. Pages 754–761 in J.D. Ralph and T. D. Rich (eds.), Bird Conservation Implementation and Integration in the Americas: Proceedings of the Third International Partners in Flight Conference, March 20–24, 2002, Asilomar, CA. Volume 2. USDA Forest Service General Technical Report PSW-GTR-191. Albany, CA: U.S. Dept. of Agriculture, Forest Service, Pacific Southwest Research Station

Orloff, S., and A. Flannery. 1992. *Wind Turbine Effects on Avian Activity, Habitat Use, and Mortality in Altamont Pass and Solano county Wind Resource Areas, 1989–1991*. P700-92-001. Report to California Energy Commission, Sacramento, CA. Santa Cruz, CA: Biosystems Analysis, Inc.

Royall, R. M., and W. G. Cumberland. 1981. An Empirical Study of the Ratio Estimator and Estimators of its Variance. *Journal of the American Statistical Association* 76:66–77.

Smallwood, K. S., and C. G. Thelander. 2004. *Developing Methods to Reduce Bird Fatalities in the Altamont Wind Resource Area*. Final Report by BioResource Consultants to the California Energy Commission, Public Interest Energy Research—Environmental Area. Contract No. 500-01-019.

Smallwood, S. 2006. Estimating Wind Turbine-Caused Bird Mortality. *Journal of Wildlife Management*. 71(8):2781–2791.

Smallwood, K. S. 2007. Estimating Wind Turbine-Caused Bird Mortality. *Journal of Wildlife Management* 71(8):2781–170.

White, G. C. 1998. *Logistic Regression Model of Detection Probability for Marbled Murrelets. Headwaters Habitat Conservation Plan.* Section IV-B(11).