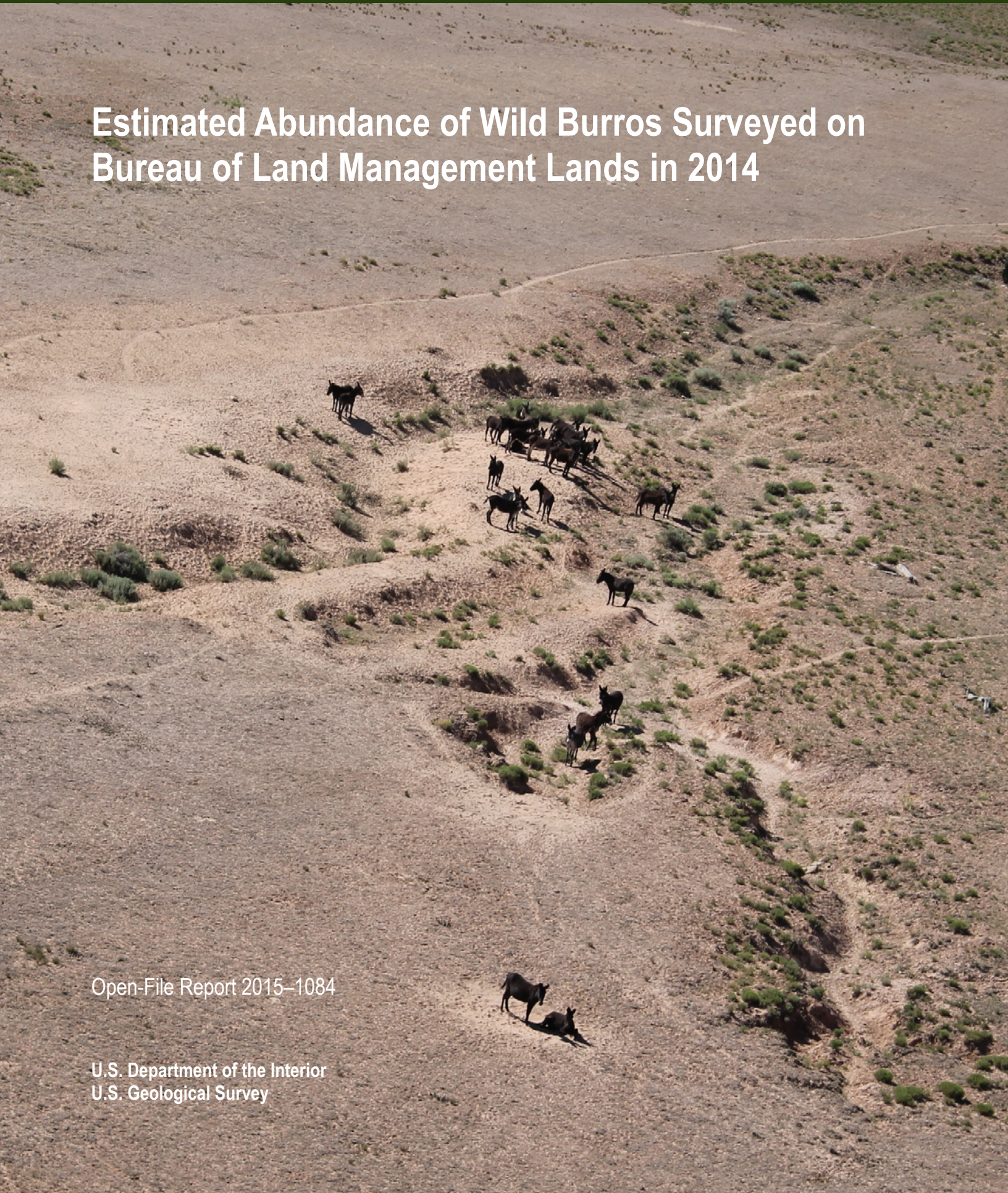


# Estimated Abundance of Wild Burros Surveyed on Bureau of Land Management Lands in 2014

Open-File Report 2015-1084

U.S. Department of the Interior  
U.S. Geological Survey



Front cover photograph: Wild burros in Sinbad Herd Management Area, Utah. Photo by D. Carpenter, Bureau of Land Management.

# **Estimated Abundance of Wild Burros Surveyed on Bureau of Land Management Lands in 2014**

By Paul C. Griffin

Open-File Report 2015–1084

**U.S. Department of the Interior  
U.S. Geological Survey**

**U.S. Department of the Interior**  
SALLY JEWELL, Secretary

**U.S. Geological Survey**  
Suzette M. Kimball, Acting Director

U.S. Geological Survey, Reston, Virginia: 2015

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment—visit <http://www.usgs.gov/> or call 1-888-ASK-USGS (1-888-275-8747).

For an overview of USGS information products, including maps, imagery, and publications, visit <http://www.usgs.gov/pubprod/>.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this information product, for the most part, is in the public domain, it also may contain copyrighted materials as noted in the text. Permission to reproduce copyrighted items must be secured from the copyright owner.

Suggested citation:

Griffin, P.C., 2015, Estimated abundance of wild burros surveyed on Bureau of Land Management Lands in 2014: U.S. Geological Survey Open-File Report 2015-1084, 42 p., <http://dx.doi.org/10.3133/ofr20151084>.

ISSN 2331-1258 (online)

# Contents

- Executive Summary..... 1
- Introduction..... 1
- Methods..... 4
  - Aerial Surveys..... 4
- Study Areas ..... 5
  - Survey Design in California, Nevada, and Utah ..... 5
  - Survey Design in Arizona..... 5
- Data Sets ..... 7
- Raw Number of Animals Counted ..... 7
- Lincoln-Petersen Double-Observer Analyses ..... 7
  - L-P 1..... 8
  - L-P 2..... 8
- Area Correction Factors ..... 9
- Huggins Double-Observer Analyses Using Sighting Covariates ..... 9
  - Huggins Model Structures ..... 10
  - Front- and Back-Seat Intercept Parameters ..... 11
  - Sighting Covariates Potentially Included in Models ..... 11
    - Group Size Effect (n) ..... 16
    - Pilot Effect (p) ..... 16
    - Side of Ship Effects ..... 16
    - Distance Effect (d) ..... 17
    - Activity Effect (a)..... 17
    - Topographic Class Effect (t) ..... 17
    - Observer Acuity Effect (o)..... 17
    - Percent Concealing Vegetation Effect (v) ..... 18
    - Visual Field Effect (f)..... 18
    - Light Effect (l) ..... 18
  - Variance Estimation for Huggins Model Analyses ..... 18
- Comparisons of Estimated Abundance ..... 19
- Results..... 19
  - Areas Surveyed ..... 19
  - Observed Burro Groups..... 19
  - Huggins Model Weights and Covariate Effects ..... 20
  - Estimated Abundance ..... 31
  - Estimated Group Sizes of Missed Groups ..... 33
- Discussion ..... 33
  - Observer Acuity ..... 36
  - Sample Size Considerations ..... 36
  - Effects of Transect Width ..... 37
- Conclusions ..... 39
- Acknowledgments ..... 39
- References Cited..... 40
- Appendix. Definitions of Sighting Covariates ..... 42

## Figures

1. Frequency distribution of recorded observation distances for the four data sets where distance was recorded in 100-meter increments.....	21
2. Distribution of distances from the Colorado River for burro group observations made in the Cibola-Trigo herd management area survey.....	21
3. Expected effect of group size on group-specific estimated overall detection probability and correction factor in six simultaneous double-observer data sets from 2014 burro aerial surveys.....	29
4. Expected effect of distance on group-specific estimated overall detection probability and correction factor in six simultaneous double-observer data sets from 2014 burro aerial surveys.....	30
5. Expected effect of the percentage of concealing vegetation on group-specific estimated overall detection probability and correction factor in three simultaneous double-observer data sets from 2014 burro aerial surveys.....	31
6. Histograms of the numbers (n) of observed burro groups and estimated numbers of missed groups in each of six herd management areas surveyed for burros in 2014.....	34

## Tables

1. Study areas where aerial surveys for wild burros took place in 2014, listed by State, Bureau of Land Management field office, and high end of the appropriate management level.....	5
2. Dates, pilots, and Federal employees participating as observers for aerial surveys for wild burros in 2014.....	6
3. Model structures considered for Huggins method analyses.....	11
4. Sample size of burro groups observed (observations) and simultaneous double-observer observations that were useful for fitting Huggins models.....	20
5. Model weights and differences in Akaike's Information Criterion scores for models fit to each data set.....	22
6. Estimated baseline detection probabilities and effects of modeled sighting covariates.....	26
7. Observed numbers of burros and estimated abundance of surveyed burro populations.....	32
8. Sighting covariates for each observed burro group.....	42

## Conversion Factors

Inch/Pound to SI

Multiply	By	To obtain
	Length	
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)

# Estimated Abundance of Wild Burros on Bureau of Land Management Lands Surveyed in 2014

By Paul C. Griffin

## Executive Summary

The Bureau of Land Management (BLM) requires accurate estimates of the numbers of wild horses (*Equus ferus caballus*) and burros (*Equus asinus*) living on the lands it manages. For over ten years, BLM in Arizona has used the simultaneous double-observer method of recording wild burros during aerial surveys and has reported population estimates for those surveys that come from two formulations of a Lincoln-Petersen type of analysis (Graham and Bell, 1989). In this report, I provide those same two types of burro population analysis for 2014 aerial survey data from six herd management areas (HMAs) in Arizona, California, Nevada, and Utah. I also provide burro population estimates based on a different form of simultaneous double-observer analysis, now in widespread use for wild horse surveys that takes into account the potential effects on detection probability of sighting covariates including group size, distance, vegetative cover, and other factors (Huggins, 1989, 1991). The true number of burros present in the six areas surveyed was not known, so population estimates made with these three types of analyses cannot be directly tested for accuracy in this report. I discuss theoretical reasons why the Huggins (1989, 1991) type of analysis should provide less biased estimates of population size than the Lincoln-Petersen analyses and why estimates from all forms of double-observer analyses are likely to be lower than the true number of animals present in the surveyed areas. I note reasons why I suggest using burro observations made at all available distances in analyses, not only those within 200 meters of the flight path. For all analytical methods, small sample sizes of observed groups can be problematic, but that sample size can be increased over time for Huggins (1989, 1991) analyses by pooling observations. I note ways by which burro population estimates could be tested for accuracy when there are radio-collared animals in the population or when there are simultaneous double-observer surveys before and after a burro gather and removal.

## Introduction

Wild burro management requires accurate estimates of the number of animals in each Bureau of Land Management (BLM) herd management area (HMA), BLM herd area (HA), and U.S. Forest Service (USFS) wild horse territory (WHT) where they occur (National Research Council, 2013). Accounting for imperfect detection probability is a priority for wild horse and burro aerial surveys conducted by BLM (BLM, 2010; National Research Council, 2013), but accurately estimating wild burro numbers in the rugged deserts and mountains of Arizona, California, Nevada, Oregon, and Utah has been an ongoing challenge to Federal and State managers. Many burro characteristics make them difficult to detect in aerial surveys. Burros are relatively small, are often cryptic-colored, and can be hidden by trees and tall shrubs. They often occur in small groups and may stand still during surveys. Past work has suggested that, depending on environmental conditions, as many as 50 percent or more of

wild burros may go unseen during a standard helicopter survey (Little and Grissom, BLM, written commun., 1999).

The BLM and Arizona Game and Fish Department (AGFD) convened the Arizona Burro Census Team, a collaborative group of specialists and staff from those agencies, to review survey methods for burros that were available in 1999. When the Arizona Burro Census Team made a review of appropriate methods for estimating burro population size, the simultaneous double-observer method (hereafter referred to as double-observer method) was identified as a relatively safe method that caused relatively little stress to the animals and could provide reasonably accurate population estimates (Little and Grissom, BLM, written commun., 1999). The double-observer method is a way of recording data in animal surveys that allows for a mark-recapture analyses of the data to estimate the number of animals that were available for observation but not seen by anyone in the helicopter. In the context of a double-observer aerial survey, the two sampling “occasions” for observing an animal group are the two (or more) independent observers, each searching for burros without giving clues to each other as to burro presence until all observers have had an opportunity to detect or miss each burro group.

Arizona BLM, in cooperation with AGFD, conducted tests of the double-observer method with pre-gather and post-gather aerial surveys in the Alamo HMA (BLM, written commun., 2000a) and Cibola-Trigo HMA (BLM, written commun., 2000b) and concluded that the survey method, with data analyzed via a Lincoln-Petersen estimator (Seber, 1982), provided population estimates that were adequately accurate and precise (BLM, written commun., 2000a, 2000b). Since that time, the BLM in Arizona has been using the double-observer method and analyzing the survey data with two forms of a Lincoln-Petersen estimator developed by Graham and Bell (1989), where the inverse of one or more observers’ average estimated detection probability is used as a correction factor that is applied to observed groups to estimate the number of animal groups present in the entire survey area.

The Lincoln-Petersen analyses used in Arizona were elaborated specifically for helicopter-based wild burro surveys as two related estimators, known as Technique 1 (L-P 1) and Technique 2 (L-P 2; Conrad and others, BLM, written commun., 2012). In both techniques, a Lincoln-Petersen analysis estimates the conditional detection probability for one or two observers, based on the number of observations made by that observer and another observer that also had opportunities to see burros on the same side of the helicopter. The main difference between the two analytical techniques is that L-P 1 uses an estimate of detection probability for the left back-seat observer, whereas L-P 2 uses an estimate of overall detection probability for the two observers on the left side of the helicopter.

There are some known drawbacks to both L-P 1 and L-P 2. Each technique relies on a single, mean estimate of detection probability. As a result, neither technique inherently accounts for group-specific differences in animal sighting probability that could result from sighting covariates, such as group size, vegetative cover, and others. Although it is true that the effects of sighting covariates influence the average detection probability, both L-P techniques may be biased by heterogeneity that is not accounted for by the average detection probability (Pollock and Kendall, 1987). In a study of feral horses, Walter and Hone (2003) showed that, if groups of animals are not all equally detectable, then the Lincoln-Peterson estimator leads to negatively biased population abundance estimates. This is because there tends to be a correlation between groups that are included in the data set and those groups that were more likely to be detected by both observers. For all double-observer analyses, the estimated detection probabilities are actually “conditional” detection probabilities because they rely only on observed burro groups that are included in the data set. As a result, the estimated conditional detection probability based on the observed data set tends to be biased high compared to true detection probability of animals in the entire population, and therefore the correction factor applied tends to be biased low, leading to an estimated population size that is lower than the true population size. The number of burros



estimated from double-observer methods should be considered only for that population of burros that could possibly have been detected by the observers; in areas with dense vegetation or extremely rugged topography, the double-observer method alone would not be expected to account for all burros in the population.

There are other clear drawbacks of L-P 1 and L-P 2. Both techniques omit from analyses any observations made by the pilot. If observing animals distracts the pilot from safely flying, then it may be reasonable to insist that the pilot not share any observations. So long as the pilot understands that his or her sole job is to fly safely, though, then any incidental observations that the pilot makes can contribute to the survey results. Also, variation in group size is not included in the variance estimation; only the mean group size of observed groups is used in the population estimate. Binomial variance in estimated detection probabilities is the only source of variation used in generating confidence intervals around population estimates. As a result, one would expect the range of L-P 1 or L-P 2 confidence intervals around estimated population sizes to be smaller than they would be if other sources of variation were considered.

The effect of heterogeneity (Southwell and others, 2007) can be reduced by developing models of sighting probability that account for effects of recorded sighting covariates on detection probabilities. Double-observer data with individual covariates for each animal group can be analyzed using a conditional likelihood approach (Huggins 1989, 1991) with logistic regression models of sighting probability (Walter and Hone, 2003). Including individual covariates into the estimation of detection probability leads to group-specific detection probability estimates and group-specific correction factors that are applied to each observed group. Relative to L-P estimators, this approach (hereafter, the Huggins method) should reduce the known positive bias in detection probabilities and associated negative bias in population estimates, due to unmodeled heterogeneity.

Even after including effects of modeled covariates, unknown and unmodeled sources of heterogeneity may cause some groups with similar covariate values to be more likely to be seen by both observers or to have been missed by both observers (Southwell and others, 2007; Barker, 2008; Laake and others, 2011). This persistent effect of heterogeneity bias can lead to an underestimate of population abundance even in Huggins analyses (Griffin and others, 2013). This “heterogeneity bias” can be modeled if an adequate sample size of radio-marked animals is present in surveyed populations (Griffin and others, 2013). Radio-marked animals can allow for a test of the adequacy of double-observer models because observers can conduct a visual survey, then immediately determine which radio-marked burros were seen or missed. A separate analytical method that might be used to account for heterogeneity bias is mark-recapture distance sampling, although this method is known to provide biased estimates if the surveyed animals move toward or away from the transect line when they hear the aircraft approaching (Burt and others, 2014).

Another way to test the accuracy of the various simultaneous double-observer analyses presented in this report would be to conduct a pre-gather aerial survey, gather and remove a substantial fraction of burros from the population, then conduct a post-gather aerial survey shortly thereafter.

U.S. Geological Survey (USGS) has a cooperative agreement to advise BLM on the design and execution of simultaneous double-observer wild horses and burro aerial surveys. As part of that agreement, USGS was tasked with analyzing the simultaneous double-observer aerial survey data for wild burro surveys conducted by BLM in 2014. The analytical techniques and results I present in this report follow the existing Lincoln-Petersen estimators used for burros in Arizona (BLM, written commun., 2000a, 2000b; Conrad and others, BLM, written commun., 2012) and Huggins methods used for wild horses on other BLM lands (Lubow and Ransom, 2007).

The objectives of this study were to:

- Compare the point estimates and precision of the L-P 1, L-P 2, and Huggins analytical methods for double-observer analysis of burro aerial survey data;
- Examine the extent to which various sighting covariates influence the apparent overall burro group detection probabilities;
- Examine the effect of truncating the maximum observation distance on L-P 1 and L-P 2 abundance estimates; and
- Make recommendations for future approaches for testing the accuracy of the above methods.

## Methods

### Aerial Surveys

All surveys analyzed in this report were conducted using a Bell 206BIII Jet Ranger helicopter or Bell 206L Long Ranger helicopter (Bell Helicopter, Fort Worth, Texas), following flight safety plans that were specifically designed for each survey. On each flight, a pilot and three experienced observers conducted surveys. The seating configuration in each survey had the pilot seated in the right front (RF) seat, and observers seated in the left front (LF) left back (LB) and right back (RB) seats. The pilot and front-seat observer had views to the front, sides, and directly below the flight path. Although pilots' attention was primarily on flying safely, in almost all surveys, pilots also detected burro groups. In the Lake Pleasant HMA survey, the pilot was encouraged to not observe burros. The pilot's only priority must always be to fly safely; so long as the pilot follows the procedures for observer independence, though, any observations that the pilot makes may contribute to the survey. Each back-seat observer had a view to one side of the flight path. The back-seat observers traded seats after refueling stops, so that each was seated behind the LF observer for some of the survey. Helicopter doors were removed for all surveys allowing for better views of the surveyed areas. In all cases, the flight lines for survey were designated ahead of time and loaded onto the pilot's global positioning system (GPS) unit to facilitate navigation and ensure full survey coverage. Flight speed was  $\approx 93\text{--}111$  kilometers per hour (km/hr) ( $\approx 50\text{--}60$  knots) at heights of  $\approx 60$  meters (m) (range: 30–100 m) above ground level, depending on topography.

I used the double-observer method for recording data (Caughley and Grice, 1982; Graham and Bell, 1989) in all surveys. The double-observer method involves recording the independent detections of available burro groups by each observer in the helicopter. In-flight procedures required each observer to search independently for burro groups in the survey area but to collaborate in determining group size, composition, and sighting covariates of detected groups after each flight member had an opportunity to detect a given burro group (for example, see Ransom, 2012). Recording data in this way allows for estimation of each observer's detection probability. I noted the rare occasions when survey protocols were not followed, such that either observer pair was inadvertently alerted about the presence of a burro group by the other before having a chance to record an observation independently, so I could omit those non-independent observations from the data sets used for Huggins method (1989, 1991) model fitting. Similarly, if one observer was entirely unavailable to observe burros (for example, due to airsickness) and the burro group was on that observer's side of the flight line, then I noted that the survey protocol was not being followed for that observation and consequently did not use the observation in model fitting.

I recorded the following covariates associated with each group: group size ( $n$ ) and composition (foals and adults; yearlings were difficult to distinguish and were counted with adults), approximate minimum distance ( $d$ ) from the preplanned transect to the burro group when it was first detected, percent concealing vegetation within 10 m of the group ( $v$ ), whether the visual field ( $f$ ) within 10 m of

the group was obscured by tall shrubs or trees, whether the group was within 10 m of rocks or cliff topography (*t*) that could have concealed a burro, the activity (*a*) of the group when first detected (moving or still), and the lighting (*l*) conditions of the group (appendix). I noted whether each burro group was detected to the left, right, or both sides of the helicopter’s flight path. I also noted those rare cases when a burro group was directly below the helicopter’s flight path and not visible to the back-seat observers (*c*; centerline yes or no). Before using these covariates in Huggins (1989, 1991) analyses, I transformed *n* by taking its natural log, converted estimated distance to the midpoint of the recorded categorical range (in kilometers), and designated a binary covariate to indicate whether the burro group was on the same side of the helicopter’s flight line as the pilot (*p*).

## Study Areas

### Survey Design in California, Nevada, and Utah

In California, Nevada, and Utah, burro surveys on BLM lands in 2014 were designed to achieve full visual coverage of the survey area. Transect spacing varied according to topography and vegetation, with wider spacing where the landscape was flat or devoid of tall vegetation. All observed burro groups were included in all analyses.

Warm Springs Canyon HMA in Nevada and High Rock HMA in California are contiguous. Because burros can move from one to the other, the two areas are managed for burros as a complex, and surveys there were conducted at the same time (table 1).

**Table 1.** Study areas where aerial surveys for wild burros took place in 2014, listed by State, Bureau of Land Management field office, and high end of the appropriate management level. Survey methods followed the simultaneous double-observer method.

[AML, appropriate management level]

Herd Management Area	State	Field Office	Appropriate management level, high
Lake Pleasant	Arizona	Hassayampa	208
Black Mountain	Arizona	Kingman	478
Cibola Trigo	Arizona	Yuma	*
McGee Mountain	Nevada	Black Rock	41
High Rock	California	Surprise	0
Warm Springs Canyon	Nevada	Black Rock	24
Sinbad	Utah	Price	70

\*The AML for Cibola Trigo is expressed as a desired population size of approximately 165 burros (BLM, 1980).

### Survey Design in Arizona

In Arizona, predetermined survey flight paths were developed following guidelines used by the BLM and AGFD since the late 1990s (Little and Grissom, BLM, written commun., 1999). Flight lines were parallel, spaced  $\approx 800$  m apart. In the past, observers only recorded burro groups less than  $\approx 200$  m from the center of the survey transect line, such that the effective total transect width combining both sides of the line was  $\approx 400$  m, meaning that only one-half of the study area was sampled.

Correspondingly, an area correction factor of two was applied to past population estimators used for Arizona burro survey data.

Generally, it is advisable to use as many of the available observations as possible in analyses, but the L-P estimator does not account for potential decreases in detection probability that come from increasing distance. Thus, the L-P estimator only accounts for distance insofar as it uses the average detection probability from all the animal groups across the observed range of distances; low-detection probability groups from greater distances are averaged in along with high-detection probability groups from closer distances. This limitation was probably the motivation behind limiting the observations to an effective strip width where detection probability was expected to be high, especially in HMAs where vegetation or topography would be expected to limit detection probabilities at longer distances.

In 2014, observers in Arizona recorded burro groups at all observable distances up to 400 m and recorded the approximate perpendicular distance from the center of the transect line to all groups. The survey crew could refer to GPS waypoints marked at the time of observations made on the previous transects to minimize the chance of recording a repeat observation of a previously-seen group. In 2014 in Arizona, at Lake Pleasant HMA and Black Mountain HMA, distances were recorded from transect to burro group in 100-m increments, whereas at Cibola Trigo HMA, distances were recorded in ¼-mile (400-m) increments.

Where possible, I used the L-P 1 and L-P 2 techniques to estimate population sizes based on only data from the 200-m effective transect widths and contrasted those estimates to similar estimates based on the entire data set of observed burro groups. Prior to 2014, observers in Arizona were only encouraged to search within 200 m of the transect line. Insofar as observers in 2014 were encouraged to search the entire area visible to them, it is conceivable that detection probability in the 0–200 m distance may have been lower than in prior years; I was unable to test that proposition directly because 2014 data were not entirely comparable to data in prior years.

The Black Mountain HMA was not completely surveyed at one time. Strong winds limited the April 2014 surveys (table 2) to the portion of the HMA north of Highway 68 and cut short some days of survey operations. The portion of the HMA south of Highway 68 was surveyed in October 2014 (table 2). In the fall surveys, the pilot and all but one of the observers was different than in the spring surveys (table 2). Because of differences in pilot and crew and potential changes in vegetation cover that could have influenced detection probabilities, I analyzed the Black Mountain HMA spring survey data separately from the fall survey data.

**Table 2.** Dates, pilots, and Federal employees participating as observers for aerial surveys for wild burros in 2014.

[Arizona Game and Fish Department (AGFD) observers identified as AGFD 1 through 9]

Herd Management Area	Dates	Pilot	Observers
Lake Pleasant	April 22–23, 2014	D. Northrup	S. Bird, R. Babel, AGFD 1
Black Mountain Spring	April 25–30, 2014	D. Northrup, J. Galloway	C. Benson, B. Peck, R. Haley, AGFD 2, AGFD 3, AGFD 4, AGFD 5, AGFD 6
Black Mountain Fall	October 6–9, 2014	D. Lyons	C. Benson, K. Birch, AGFD 7, AGFD 8
Cibola Trigo	May 2–3, 2014	J. Galloway	J. Hall, C. Weise, AGFD 9
McGee Mountain	July 15, 2014	J. Kelly	G. Swisher, S. Surian, P. Griffin
High Rock	July 16, 2014	J. Kelly	G. Swisher, S. Surian, P. Griffin
Warm Springs Canyon	July 16–17, 2014	J. Kelly	G. Swisher, S. Surian, P. Griffin
Sinbad	June 26, 2014	J. Martin	M. Tweddell, D. Carpenter, P. Griffin

Three additional HMAs were surveyed for burros in 2014, but data from those surveys are not analyzed in this report. Hickison Summit HMA in Nevada was surveyed using simultaneous double-observer methods in March 2014 along with Hickison Summit WHT on USFS lands. Marietta Wild Burro Range was surveyed using simultaneous double-observer methods in August 2014. Burro populations in those areas will be estimated at a later date. Centennial HMA, located on Department of Defense lands in California, was surveyed with an incomplete direct count, so there will be no double-observer analysis of those data.

## **Data Sets**

I compiled detection histories for all observed burro groups, denoting exactly which observer or observers detected each burro group in the data set. A burro group was considered to be any group of one or more burros that move together as part of the same visually identifiable social unit at the time of detection and which get detected at more or less the same time in the process of observers searching for burros.

I considered pooling double-observer data from different HMAs when there was a plausible expectation that estimated detection probabilities may be comparable from those data sets. Differences in individual observer acuity can lead to large differences in estimated detection probabilities.

As the observers and pilot were the same for all three days of flights over McGee Mountain HMA, Warm Springs Canyon HMA, and High Rock HMA (table 2), I analyzed the data set from all three together, though I produced separate abundance estimates for each HMA. Some burros were observed during fixed-wing surveys over these HMAs as part of the June 1–11, 2014, “tri-State” surveys for horses, but I did not use any of those data in the analyses presented here because wild burro detection probability from fixed-wing surveys is generally considered too low to provide reliable data. Data sets from all other HMAs were analyzed separately because few or none of the observers and pilots overlapped.

## **Raw Number of Animals Counted**

The uncorrected tally of animals counted during surveys can be considered to be the minimum number of animals (MNA) known to be alive at the time of the survey. MNA is the index to abundance that comes from direct-count aerial surveys; it is the raw number of animals counted. The MNA index by itself is problematic because it has neither a known relation to true abundance (Anderson, 2001) nor any associated uncertainty estimates.

In contrast to direct counts, the simultaneous double-observer method for recording data can be used to calculate a population estimate of the total number of animals present in the survey area, including those that were not seen by any of the observers, and of confidence intervals around that population estimate.

## **Lincoln-Petersen Double-Observer Analyses**

The Lincoln-Petersen analyses included two techniques for estimating abundance and associated confidence intervals, L-P 1 and L-P 2. These techniques essentially differ only in terms of which observer’s or observers’ estimated detection probability is used to correct for the number of animals that are not seen by any observer. In this report, I followed the existing analytical methods used by BLM and AGFD for burro population estimation in recent years, using both of these techniques and observations limited to those <200 m from the transect (Conrad and others, BLM, written commun., 2012); I also explored the application of the same techniques using all available observations.

Observed burro groups used in the data set are classified as having been detected by only the LF observer, only the LB observer, only the RB observer, or both LF and LB observers. The total number

of such groups in each data set are termed *S1*, *S2*, *S3*, and *B*, respectively, in analyses. Burro groups observed beyond the detection distance limit, which was historically  $\approx 200$  m, were used in calculations of average group size of observed burro groups but not in estimates of detection probability or abundance. Observations made only by the pilot are not included in the L-P 1 or L-P 2 analyses.

## L-P 1

The correction factor used in this technique is based on the estimated L-P burro group detection probability for the left back-seat observer ( $\hat{p}_{LB,L-P}$ , also called *P2*). A correction factor based on  $\hat{p}_{LB,L-P}$  is applied to all groups observed by back-seat observers, reflecting an assumption that the left back-seat observer's detection probability is equal to that of the right back-seat observer. The left back-seat observer's probability is estimated based only on observations made by the LF or LB observers, according to the equation:

$$\hat{p}_{LB,L-P} = P2 = B/(S1 + B)$$

Here, *B* is the number of burro groups seen by both the LF and LB observers, and *S1* is the number of burro groups seen only by the LF observer. Variance for  $\hat{p}_{LB,L-P}$ , also called *V2*, is based on the binomial distribution and decreases as a function of the sample size of groups used to estimate  $\hat{p}_{LB,L-P}$ .

$$Var(\hat{p}_{LB,L-P}) = V2 = \hat{p}_{LB,L-P} * \frac{(1 - \hat{p}_{LB,L-P})}{(S1 + B)}$$

To calculate the population estimate for L-P 1 ( $\hat{N}_{L-P 1}$ ), the total number of groups seen by the LB and RB observers within the designated transect areas is multiplied by the average group size of all observed groups ( $\bar{n}_{observed}$ ) and the area correction factor (*A*), and divided by  $\hat{p}_{LB,L-P}$ .

$$\hat{N}_{L-P 1} = \frac{(S2 + S3 + B) * \bar{n}_{observed} * A}{\hat{p}_{LB,L-P}}$$

The area correction factor (*A*) is the inverse of the proportion of the sampling frame that was surveyed. In analyses of Arizona data where the burro groups used are limited to those within  $\approx 200$  m of the transect line but where transects are spaced at  $\approx 800$  m, *A* = 2. For analyses where all burro groups in the data set are used, and the survey is considered to have covered the sample frame completely, *A* = 1.

The confidence intervals (CI) around the population abundance estimate for L-P 1 only incorporate the estimated variance for  $\hat{p}_{LB,L-P}$ , and do not include any variance term for observed group size.

$$90\% CI (\hat{N}_{L-P 1}) = \frac{(S2 + S3 + B) * \bar{n}_{observed} * A}{\hat{p}_{LB,L-P} \pm T * SE(\hat{p}_{LB,L-P})}$$

In the equation above, the standard error of  $\hat{p}_{LB,L-P}$ ,  $SE(\hat{p}_{LB,L-P})$ , is the square root of  $Var(\hat{p}_{LB,L-P})$ , and *T* is the two-tailed value for a t-distribution, given an associated Type I error (for 90 percent confidence interval, two-tailed  $\alpha = 0.1$ ) and where degrees of freedom (*df*) is the number of groups seen by the LF observer (*S1* + *B*) minus 1.

## L-P 2

The correction factor used in L-P 2 is based on the estimated overall L-P detection probability for both of the left-seat observers ( $\hat{p}_{Left,L-P}^*$ ), also sometimes called *p<sub>both</sub>*, which leads to a correction factor that is applied to all burro groups observed on the left side of the transect. This analysis doubles the estimated number of burro groups on the left side of the transect to make inferences to the total population of burros on both sides of the transect. The L-P 2 analysis may be considered a technique to overcome the low number of observations to be expected from the pilot's side (right side) of the

helicopter on surveys when the pilot is encouraged to not make any observations. The overall left-seat observers' detection probability is calculated based on the estimated L-P detection probability for the front left observer ( $\hat{p}_{LF,L-P}$ ) and for the back left observer ( $\hat{p}_{LB,L-P}$ ). The latter is calculated as described for technique L-P 1, but  $\hat{p}_{LF,L-P}$  is calculated as

$$\hat{p}_{LF,L-P} = S1 = B/(S2 + B)$$

Here, S2 is the number of burro groups detected only by the LB observer. Overall detection probability for left side observers is one minus the probability that both LF and LB observers missed:

$$\hat{p}^*_{Left,L-P} = P_{both} = 1 - (1 - \hat{p}_{LF,L-P}) * (1 - \hat{p}_{LB,L-P})$$

The estimator used for the variance of  $\hat{p}^*_{Left,L-P}$ , also sometimes called  $V_{both}$ , has been

$$Var(\hat{p}^*_{Left,L-P}) = V_{both} = \frac{S1 * S2 * (S1 + S2) * \hat{p}^*_{Left,L-P}}{(S1 + B)^2 * (S2 + B)^2}$$

The estimated population size calculated in technique L-P 2 is similar in form to that for L-P 1, except that the number of groups used in the numerator includes all groups seen by any observer on the left side and does not include any groups seen by the RB observer. In addition to the area correction factor,  $A$ , that potentially accounts for effective strip width, the numerator is multiplied by 2 to reflect the assumption that there are an equal number of burros on the right side of the transect line as there are on the left side of the transect line.

$$\hat{N}_{L-P 2} = \frac{(S1 + S2 + B) * \bar{n}_{observed} * A * 2}{\hat{p}^*_{Left,L-P}}$$

The confidence intervals around  $\hat{N}_{L-P 2}$  only incorporate the estimated variance for  $\hat{p}^*_{Left,L-P}$ , and do not include any variance of observed group size.

$$90\% CI (\hat{N}_{L-P 2}) = \frac{(S1 + S2 + B) * \bar{n}_{observed} * A * 2}{\hat{p}^*_{Left,L-P} \pm T * SE(\hat{p}^*_{Left,L-P})}$$

When determining the value of  $T$  from the t-distribution, given  $\alpha$ ,  $df$  is the total number of groups seen by the LB observer ( $S2 + B$ ) minus 1.

### Area Correction Factors

In Arizona in the past, the systematic survey on 800 m spacing and maximum search distance of 200 m led to only half of the sampling frame being surveyed, so that  $A$  was 2 for those surveys. In 2014 in Arizona, I estimated L-P 1 and L-P 2 population abundance with reference to two possible data sets and associated area correction factors. First, I used all observations made during the survey and applied an area correction factor of one. Second, I used only those observations that were recorded  $\leq 200$  m from the transect, as was done in past surveys, and applied an area correction factor of 2 in the resulting L-P 1 and L-P 2 estimates.

In California, Nevada, and Utah in 2014, the surveys were designed to cover the entire sample frame, and transect width varied according to topography and vegetation, so  $A$  was one.

### Huggins Double-Observer Analyses Using Sighting Covariates

The effect of heterogeneity (Southwell and others, 2007) can be reduced by developing models of sighting probability that account for effects of recorded sighting covariates on detection probabilities. Double-observer data with individual covariates for each animal group can be analyzed using a conditional likelihood approach (Huggins, 1989, 1991) with logistic regression models of sighting probability (Walter and Hone, 2003). Including individual covariates into the estimation of detection probability leads to group-specific correction factors. Relative to L-P estimators, this approach should

reduce the known positive bias in detection probabilities and associated negative bias in correction factors coming from unmodeled heterogeneity. Even after including effects of modeled covariates, unmodeled sources of heterogeneity can cause some groups with similar covariate values to be more likely to be seen by both observers or missed by both observers (Southwell and others, 2007; Barker, 2008; Laake and others, 2011). This “heterogeneity bias” can be modeled if an adequate sample size of radio-marked animals is present in surveyed populations (Griffin and others, 2013). For these surveyed populations of burros, though, no radio-collared animals were present.

### Huggins Model Structures

My general strategy for determining which sighting covariate effects influenced detection probability in a given data set was to develop a candidate suite (Burnham and Anderson, 2002) of Huggins model structures (table 3), each of which was a logistic regression model with a logit link that included parameters to represent a number of potential effects (Huggins, 1989, 1991). I fit the candidate models to the data using program MARK (White and Burnham, 1999). I used Akaike’s Information Criterion scores adjusted for small sample size (AICc) to assess the parsimony of each  $k$  model and used model-averaging to apply associated model weights ( $w_k$ ) to estimates coming from each  $k$  model (Burnham and Anderson, 2002).

If either the pilot or the front-seat observer detected a group, I considered the group as having been detected by the front observer pair. If either of the back-seat observers detected a group, I considered the group as having been detected by the back observer pair.

For each survey data set, I defined sighting probabilities,  $\hat{p}_{h,j,k}$ , where  $h$  is the observer ( $F$  = front, or  $B$  = back),  $j$  indexes the burro group (1, 2, . . . ,  $n$ , where  $n$  is the number of groups detected), and  $k$  indexes the hypothesized alternative model parameterization of each type ( $k = 1, 2, . . . , M$ , where  $M$  is the number of model parameterizations for that data set).

I modeled sighting probabilities for observer  $h$ , burro group  $j$ , according to model structure  $k$  by the logistic function (Huggins, 1989, 1991):

$$\hat{p}_{h,j,k} = \frac{e^{(\sum_i \beta_{k,i} * D_{h,k,i} * x_{i,j})}}{1 + e^{(\sum_i \beta_{k,i} * D_{h,k,i} * x_{i,j})}}$$

for { $h$ : F, B}; { $j$ : 1, 2 . . .  $n$ }; { $k$ : 1, , 2 . . .  $m$ },

where

$\hat{\beta}_{k,i}$  = is a fitted coefficient for model  $k$ , where  $i = 1, 2, . . . p$  indexes the list of all  $p$  candidate coefficients in all models.

$D_{h,k,i}$  = is an indicator for model  $k$  specific observer  $h$ . The value is set to 1 if parameter  $i$  is included in model  $k$ , and 0 otherwise.

$x_{i,j}$  = is the observed value of sighting covariate  $i$  for burro group  $j$ .

For each candidate model,  $k$ , I computed the expected unconditional overall probability that  $\geq 1$  observer would detect any  $j$  given group,  $\hat{p}_{\bullet,j,k}$ , (dot notation for the position index indicates overall sighting probability for all observers) as 1 minus the expected probability that it was missed by all observers:

$$\hat{p}_{\bullet,j,k} = 1 - (1 - \hat{p}_{F,j,k}) * (1 - \hat{p}_{B,j,k})$$

For each of the  $k$  models, I computed a group-specific correction factor,  $\hat{\theta}_{j,k}$ , for each  $j$  group, as  $1/\hat{p}_{\bullet,j,k}$ . The application of such Horvitz-Thompson correction factors for estimating detection biases in aerial surveys is well established (Steinhorst and Samuel, 1989; Borchers and others, 1998). Group-specific, model-specific, correction factors derived from  $\hat{p}_{\bullet,j,k}$  can be thought of as the per-observed-burro contribution to the overall estimate of abundance, according to model  $k$ . For each of the  $k$  models



considered, the number of burros that each observed group represents in the estimated total population,  $\bar{N}$ , is the product of group size,  $n_j$ , and the group-specific correction factor,  $n_j \hat{\theta}_{j,k}$ . These estimates are weighted by model weight,  $\hat{w}_k$ , and summed across all observed groups to calculate the total estimated abundance. Thus, I computed a model-averaged estimate of abundance:

$$\bar{N} = \sum_j \sum_k n_j \hat{\theta}_{j,k} * \hat{w}_k$$

Each  $j$  observed group represented  $\sum_k (\hat{\theta}_{j,k} - 1) * \hat{w}_k$  missed groups. As a result, the observed  $n_j$  burros in each  $j$  group represented  $\hat{n}_{j,misssed}$  groups, where

$$\hat{n}_{j,misssed} = \sum_k n_j (\hat{\theta}_{j,k} - 1) * \hat{w}_k$$

For each survey, I estimated the ratio of foals to adults in the population from the total estimated numbers of foals and adults by multiplying the observed numbers of foals and adults in each group,  $j$ , by the group-specific correction factor, such that the estimated total number of burros in age-sex class  $c$  (foal or adult),  $\hat{N}_c$ , was

$$\hat{N}_c = \sum_j \sum_k n_{j,c} \hat{\theta}_{j,k} * \hat{w}_k$$

Here, and  $n_{j,c}$  is the number of burros of age-sex class  $c$  seen in group  $j$ .

### Front- and Back-Seat Intercept Parameters

Within the candidate suite of models fit to each data set (table 3), some models had front and back observer intercept parameters set to be equal ( $f=b$ ), and some allowed the front- and back-seat intercepts to be unique ( $f \neq b$ ). These two underlying structures reflect the fact that I relied on the data to indicate whether front- and back-seat detection probabilities would be equal or different in each data set.

### Sighting Covariates Potentially Included in Models

If there were  $\leq 10$  observations with any variation in a given sighting covariate in a data set, then the model with that covariate effect on detection probability was not considered. The potential effects that were either included or excluded from consideration in the model suite fit to each data set are shown in table 3 and described below. The group size effect and pilot's side effect were included in all models.

**Table 3.** Model structures considered for Huggins method analyses. Each model is listed with a model number, indication of whether intercepts were modeled as equal for the front- and back-seat observers ( $f=b$ ) or as unequal ( $f \neq b$ ), and a list of covariates included as additive effects in the model for detection probability. All models included a covariate for the effect on front-seat detection probability of a group being only on the pilot's side of ship ( $p$ ) and one covariate for the natural log of group size ( $n$ ). For data sets with adequate variation, other covariates potentially included in models were activity ( $a$ ), percent concealing vegetation ( $v$ ), visual field ( $f$ ), distance ( $d$ ), observer acuity ( $o$ ), topographic class ( $t$ ), and light ( $l$ ). The number of parameters in each model is listed as  $K$ .

Model number	Observer intercept	Covariate effects included	K	Lake Pleasant	Cibola Trigo <sup>1</sup>	Black Mountain, spring	Black Mountain, fall <sup>1</sup>	Bi-State <sup>2</sup>	Sinbad
M1	$f=b$	$p, n$	3	X	X	X	X	X	X
M2	$f \neq b$	$p, n, a$	4		X	X	X		X
M3	$f=b$	$p, n, a, v$	5			X	X		

**Table 3.** Model structures considered for Huggins method analyses. Each model is listed with a model number, indication of whether intercepts were modeled as equal for the front- and back-seat observers (f=b) or as unequal (f≠b), and a list of covariates included as additive effects in the model for detection probability. All models included a covariate for the effect on front-seat detection probability of a group being only on the pilot's side of ship (p) and one covariate for the natural log of group size (n). For data sets with adequate variation, other covariates potentially included in models were activity (a), percent concealing vegetation (v), visual field (f), distance (d), observer acuity (o), topographic class (t), and light (l). The number of parameters in each model is listed as K.—Continued

Model number	Observer intercept	Covariate effects included	K	Lake Pleasant	Cibola Trigo <sup>1</sup>	Black Mountain, spring	Black Mountain, fall <sup>1</sup>	Bi-State <sup>2</sup>	Sinbad
M4	f=b	p, n, a, v, d	6			X	X		
M5	f=b	p, n, a, d	5		X	X	X		X
M6	f=b	p, n, v	4	X		X	X		
M7	f=b	p, n, d	4	X	X	X	X	X	X
M8	f=b	p, n, v, d	5	X		X	X		
M9	f≠b	p, n	4	X	X	X	X	X	X
M10	f≠b	p, n, a	5		X	X	X		X
M11	f≠b	p, n, a, v	6			X	X		
M12	f≠b	p, n, a, v, d	7			X	X		
M13	f≠b	p, n, a, d	6		X	X	X		X
M14	f≠b	p, n, v	5	X		X	X		
M15	f≠b	p, n, d	5	X	X	X	X	X	X
M16	f≠b	p, n, v, d	6	X		X	X		
M17	f=b	p, n, o	3	X	X	X	X	X	X
M18	f=b	p, n, a, o	4		X	X	X		X
M19	f=b	p, n, a, v, o	5			X	X		
M20	f=b	p, n, a, v, d, o	6			X	X		
M21	f=b	p, n, a, d, o	5		X	X	X		X
M22	f=b	p, n, v, o	4	X		X	X		
M23	f=b	p, n, d, o	4	X	X	X	X	X	X
M24	f=b	p, n, v, d, o	5	X		X	X		
M25	f≠b	p, n, o	4	X	X	X	X	X	X
M26	f≠b	p, n, a, o	5		X	X	X		X
M27	f≠b	p, n, a, v, o	6			X	X		
M28	f≠b	p, n, a, v, d, o	7			X	X		
M29	f≠b	p, n, a, d, o	6		X	X	X		X
M30	f≠b	p, n, v, o	5	X		X	X		
M31	f≠b	p, n, d, o	5	X	X	X	X	X	X
M32	f≠b	p, n, v, d, o	6	X		X	X		
M33	f=b	p, n, a, f	5		X	X	X		
M34	f=b	p, n, a, f, d	6		X	X	X		
M35	f=b	p, n, f	4	X	X	X	X		
M36	f=b	p, n, f, d	5	X	X	X	X		
M37	f≠b	p, n, a, f	6		X	X	X		

**Table 3.** Model structures considered for Huggins method analyses. Each model is listed with a model number, indication of whether intercepts were modeled as equal for the front- and back-seat observers (f=b) or as unequal (f≠b), and a list of covariates included as additive effects in the model for detection probability. All models included a covariate for the effect on front-seat detection probability of a group being only on the pilot's side of ship (p) and one covariate for the natural log of group size (n). For data sets with adequate variation, other covariates potentially included in models were activity (a), percent concealing vegetation (v), visual field (f), distance (d), observer acuity (o), topographic class (t), and light (l). The number of parameters in each model is listed as K.—Continued

Model number	Observer intercept	Covariate effects included	K	Lake Pleasant	Cibola Trigo <sup>1</sup>	Black Mountain, spring	Black Mountain, fall <sup>1</sup>	Bi-State <sup>2</sup>	Sinbad
M38	f≠b	p, n, a, f, d	7		X	X	X		
M39	f≠b	p, n, f	5	X	X	X	X		
M40	f≠b	p, n, f, d	6	X	X	X	X		
M41	f=b	p, n, a, f, o	6		X	X	X		
M42	f=b	p, n, a, f, d, o	7		X	X	X		
M43	f=b	p, n, f, o	5	X	X	X	X		
M44	f=b	p, n, f, d, o	6	X	X	X	X		
M45	f≠b	p, n, a, f, o	7		X	X	X		
M46	f≠b	p, n, a, f, d, o	8		X	X	X		
M47	f≠b	p, n, f, o	5	X	X	X	X		
M48	f≠b	p, n, f, d, o	6	X	X	X	X		
M49	f=b	p, n, t	4		X	X	X		
M50	f=b	p, n, a, t	5		X	X	X		
M51	f=b	p, n, a, v, t	6			X	X		
M52	f=b	p, n, a, v, d, t	7			X	X		
M53	f=b	p, n, a, d, t	6		X	X	X		
M54	f=b	p, n, v, t	5			X	X		
M55	f=b	p, n, d, t	5		X	X	X		
M56	f=b	p, n, v, d, t	6			X	X		
M57	f≠b	p, n, t	5		X	X	X		
M58	f≠b	p, n, a, t	6		X	X	X		
M59	f≠b	p, n, a, v, t	7			X	X		
M60	f≠b	p, n, a, v, d, t	8			X	X		
M61	f≠b	p, n, a, d, t	7		X	X	X		
M62	f≠b	p, n, v, t	6			X	X		
M63	f≠b	p, n, d, t	6		X	X	X		
M64	f≠b	p, n, v, d, t	7			X	X		
M65	f=b	p, n, o, t	4		X	X	X		
M66	f=b	p, n, a, o, t	5		X	X	X		
M67	f=b	p, n, a, v, o, t	6			X	X		
M68	f=b	p, n, a, v, d, o, t	7			X	X		
M69	f=b	p, n, a, d, o, t	6		X	X	X		
M70	f=b	p, n, v, o, t	5			X	X		
M71	f=b	p, n, d, o, t	5		X	X	X		
M72	f=b	p, n, v, d, o, t	6			X	X		

**Table 3.** Model structures considered for Huggins method analyses. Each model is listed with a model number, indication of whether intercepts were modeled as equal for the front- and back-seat observers (f=b) or as unequal (f≠b), and a list of covariates included as additive effects in the model for detection probability. All models included a covariate for the effect on front-seat detection probability of a group being only on the pilot's side of ship (p) and one covariate for the natural log of group size (n). For data sets with adequate variation, other covariates potentially included in models were activity (a), percent concealing vegetation (v), visual field (f), distance (d), observer acuity (o), topographic class (t), and light (l). The number of parameters in each model is listed as K.—Continued

Model number	Observer intercept	Covariate effects included	K	Lake Pleasant	Cibola Trigo <sup>1</sup>	Black Mountain, spring	Black Mountain, fall <sup>1</sup>	Bi-State <sup>2</sup>	Sinbad
M73	f=b	p, n, o, t	5		X	X	X		
M74	f=b	p, n, a, o, t	6		X	X	X		
M75	f=b	p, n, a, v, o, t	7			X	X		
M76	f=b	p, n, a, v, d, o, t	8			X	X		
M77	f=b	p, n, a, d, o, t	7		X	X	X		
M78	f=b	p, n, v, o, t	6			X	X		
M79	f=b	p, n, d, o, t	6		X	X	X		
M80	f=b	p, n, v, d, o, t	7			X	X		
M81	f=b	p, n, a, f, t	6		X	X	X		
M82	f=b	p, n, a, f, d, t	7		X	X	X		
M83	f=b	p, n, f, t	5		X	X	X		
M84	f=b	p, n, f, d, t	6		X	X	X		
M85	f=b	p, n, a, f, t	7		X	X	X		
M86	f=b	p, n, a, f, d, t	8		X	X	X		
M87	f=b	p, n, f, t	6		X	X	X		
M88	f=b	p, n, f, d, t	7		X	X	X		
M89	f=b	p, n, a, f, o, t	7		X	X	X		
M90	f=b	p, n, a, f, d, o, t	8		X	X	X		
M91	f=b	p, n, f, o, t	6		X	X	X		
M92	f=b	p, n, f, d, o, t	7		X	X	X		
M93	f=b	p, n, a, f, o, t	8		X	X	X		
M94	f=b	p, n, a, f, d, o, t	9		X	X	X		
M95	f=b	p, n, f, o, t	6		X	X	X		
M96	f=b	p, n, f, d, o, t	7		X	X	X		
M97	f=b	p, n, l	5				X		
M98	f=b	p, n, a, l	6				X		
M99	f=b	p, n, a, v, l	7				X		
M100	f=b	p, n, a, v, d, l	8				X		
M101	f=b	p, n, a, d, l	7				X		
M102	f=b	p, n, v, l	6				X		
M103	f=b	p, n, d, l	6				X		
M104	f=b	p, n, v, d, l	7				X		
M105	f=b	p, n, o, l	5				X		
M106	f=b	p, n, a, o, l	6				X		

**Table 3.** Model structures considered for Huggins method analyses. Each model is listed with a model number, indication of whether intercepts were modeled as equal for the front- and back-seat observers (f=b) or as unequal (f≠b), and a list of covariates included as additive effects in the model for detection probability. All models included a covariate for the effect on front-seat detection probability of a group being only on the pilot's side of ship (p) and one covariate for the natural log of group size (n). For data sets with adequate variation, other covariates potentially included in models were activity (a), percent concealing vegetation (v), visual field (f), distance (d), observer acuity (o), topographic class (t), and light (l). The number of parameters in each model is listed as K.—Continued

Model number	Observer intercept	Covariate effects included	K	Lake Pleasant	Cibola Trigo <sup>1</sup>	Black Mountain, spring	Black Mountain, fall <sup>1</sup>	Bi-State <sup>2</sup>	Sinbad
M107	f=b	p, n, a, v, o, l	7				X		
M108	f=b	p, n, a, v, d, o, l	8				X		
M109	f=b	p, n, a, d, o, l	7				X		
M110	f=b	p, n, v, o, l	6				X		
M111	f=b	p, n, d, o, l	6				X		
M112	f=b	p, n, v, d, o, l	7				X		
M113	f=b	p, n, a, f, l	7				X		
M114	f=b	p, n, a, f, d, l	8				X		
M115	f=b	p, n, f, l	6				X		
M116	f=b	p, n, f, d, l	7				X		
M117	f=b	p, n, a, f, o, l	8				X		
M118	f=b	p, n, a, f, d, o, l	9				X		
M119	f=b	p, n, f, o, l	6				X		
M120	f=b	p, n, f, d, o, l	7				X		
M121	f=b	p, n, t, l	6				X		
M122	f=b	p, n, a, t, l	7				X		
M123	f=b	p, n, a, v, t, l	8				X		
M124	f=b	p, n, a, v, d, t, l	9				X		
M125	f=b	p, n, a, d, t, l	8				X		
M126	f=b	p, n, v, t, l	7				X		
M127	f=b	p, n, d, t, l	7				X		
M128	f=b	p, n, v, d, t, l	8				X		
M129	f=b	p, n, o, t, l	6				X		
M130	f=b	p, n, a, o, t, l	7				X		
M131	f=b	p, n, a, v, o, t, l	8				X		
M132	f=b	p, n, a, v, d, o, t, l	9				X		
M133	f=b	p, n, a, d, o, t, l	8				X		
M134	f=b	p, n, v, o, t, l	7				X		
M135	f=b	p, n, d, o, t, l	7				X		
M136	f=b	p, n, v, d, o, t, l	8				X		
M137	f=b	p, n, a, f, t, l	8				X		
M138	f=b	p, n, a, f, d, t, l	9				X		
M139	f=b	p, n, f, t, l	7				X		
M140	f=b	p, n, f, d, t, l	8				X		
M141	f=b	p, n, a, f, o, t, l	9				X		

**Table 3.** Model structures considered for Huggins method analyses. Each model is listed with a model number, indication of whether intercepts were modeled as equal for the front- and back-seat observers (f=b) or as unequal (f≠b), and a list of covariates included as additive effects in the model for detection probability. All models included a covariate for the effect on front-seat detection probability of a group being only on the pilot's side of ship (p) and one covariate for the natural log of group size (n). For data sets with adequate variation, other covariates potentially included in models were activity (a), percent concealing vegetation (v), visual field (f), distance (d), observer acuity (o), topographic class (t), and light (l). The number of parameters in each model is listed as K.—Continued

Model number	Observer intercept	Covariate effects included	K	Lake Pleasant	Cibola Trigo <sup>1</sup>	Black Mountain, spring	Black Mountain, fall <sup>1</sup>	Bi-State <sup>2</sup>	Sinbad
M142	f=b	p, n, a, f, d, o, t, l	10				X		
M143	f≠b	p, n, f, o, t, l	7				X		
M144	f≠b	p, n, f, d, o, t, l	8				X		

<sup>1</sup>For Cibola Trigo herd management area (HMA) and Black Mountain HMA fall surveys, the observer effect included effects of two nonintercept back-seat observers, so the number of parameters in these models was the listed value of K, plus 1.

<sup>2</sup>The Bi-State surveys included observations from McGee Mountain HMA and Warm Springs Canyon HMA.

### Group Size Effect (n)

All models included an effect of group size. The covariate used in modeling was the natural log of the total group size, where total group size was the sum of adults and foals. This effect was included in all models considered because it is well established that larger group sizes tend to have higher detection probabilities than smaller group sizes. For one observation at Warm Springs Canyon HMA where one burro was seen with four horses, the group size used for the group's detection probability and associated correction factor was five, although only the single burro from the group was considered in estimates of population size.

### Pilot Effect (p)

I used an additive parameter to model front-seat detection probability for burro groups that were recorded as being only on the same side of the helicopter's flight path as the pilot's seat (the right side). I expected that the pilot's attention would be focused on safe flight and that any burros exclusively on the pilot's side of the flight path could have potentially different front-seat detection probability.

### Side of Ship Effects

In the Black Mountain HMA fall 2014 data set, there were 16 observations along the centerline of the flight path. In the Sinbad HMA data set, there was one observation recorded along the centerline of the flight path. So, for all models fit to data from both Sinbad HMA and Black Mountain fall 2014, I also included a potential effect of the observation being along the centerline on rear-seat detection probability. In these analyses,  $\hat{p}_{B,j,k}$  was fixed to equal 0 for any group along the centerline.

For the Cibola-Trigo data set,  $\hat{p}_{B,j,k}$  was fixed to equal 1 for the three observed groups recorded as being on both sides of the flight line, as they were all observed by both back-seat observers. In the Black Mountain HMA fall 2014 data set, not all of the 21 observations recorded as being visible on both sides of the flight line were observed. In the logit links for both front- and back-seat observers in models fit to that data set, I included an additive parameter for observations available to observers on both sides of the helicopter.

#### Distance Effect (d)

I included perpendicular distance from the transect line as a potential sighting covariate. There was adequate variation in all data sets to consider distance a potential effect to include in candidate model suites. Distance was recorded in ¼-mile (400-m) increments at Cibola Trigo HMA and Sinbad HMA, but in 100-m increments at other HMAs. Distance was modeled as the midpoint of the distance range circled in the field. Thus, recorded distances of <100 m, 100–200 m, 200–300 m, 300–400 m, <¼ mile, and ¼–½ mile were converted for modeling to distances, in kilometers, of 0.050 km, 0.150 km, 0.250 km, 0.350 km, 0.220 km, and 0.660 km, respectively. The distance used for two observations recorded at Lake Pleasant HMA as >400 m were 0.440 km and 0.550 km, based on values noted in the comments section of the data sheet.

For cases where I suspected that the distance measure was too large, relative to the spacing between lines, I used geographic information system (GIS) to examine the flight path of the helicopter and the location where the helicopter circled around the burro group. In analyses, I used the smaller of two distances: either the approximate distance recorded in flight or the measured perpendicular distance from flight transect line to centroid of the orbit around the burro group.

#### Activity Effect (a)

Burros were recorded in the field as either “still” or “moving.” Still was considered to be the intercept, and the effect of moving was modeled an additive parameter. The effect of activity is always expected to be positive because groups of moving animals tend to be easier to detect.

In the data set from McGee Mountain HMA and Warm Springs Canyon HMA, initial model fitting indicated that the effect of movement was opposite to that expected; in models that included that effect, moving animals were expected to be less likely to have been detected. That data set had a small sample size of observations in which to fit models (n=25), so I considered the estimated effect to be spurious, and I omitted the effect of movement from model structures considered with that data set.

#### Topographic Class Effect (t)

For data sets in which where there was adequate variation in the topographic class covariate, I modeled “smooth” as the intercept and “rugged” as an additive parameter. There was not enough variation in the McGee HMA/Warm Springs Canyon HMA/High Rock HMA data set or in the Sinbad HMA data set to include topographic class as a potential effect in model structures considered.

#### Observer Acuity Effect (o)

For data sets where there was an adequate number of observations to test for an effect of observer acuity, I structured some models to include an effect of a particular back-seat observer individual on back-seat detection probability. In the Lake Pleasant HMA, Sinbad HMA, and McGee Mountain HMA/Warm Springs HMA/High Rock Canyon HMA data sets there were only two back-seat observers in the surveys, so one back-seat observer’s detection probabilities were modeled as the back-seat intercept, and the effect of the other observer’s acuity was potentially modeled as an additive parameter (or parameters). In the Cibola Trigo HMA and Black Mountain HMA spring survey data sets there were more than two nonintercept back-seat observers. For Cibola Trigo, the observer with the greatest number of potential observations (observer JH) was considered the intercept for detection probability in the logit link, and the other two observers were each modeled as additive parameters, both of which were included in every model that had an observer effect in the structure. For Black Mountain HMA spring surveys, there were so many back-seat observers—most of which had few possible observations—that I modeled a single intercept for all of them, and the one back-seat observer who was

in the helicopter for all flights (observer CB) was modeled as an additive effect in models that included observer effects in the structures.

#### Percent Concealing Vegetation Effect (v)

There was adequate variation in the Lake Pleasant HMA and Black Mountain HMA data sets to consider models that included an effect of percent concealing vegetation. The covariate was recorded in 10-percent increments and was calculated in models as a decimal fraction from 0 to 1.0.

#### Visual Field Effect (f)

There was not enough variation in the visual field covariate to include it in any models for the McGee HMA or Sinbad HMA data sets. For the Lake Pleasant HMA, Cibola Trigo HMA, and Black Mountain HMA data sets, I analyzed the effect of visual field by using observations where the recorded visual field was “open” as the intercept and anything but open (“broken” or “tree”) as an additive parameter in the logit link. I never included the effect of visual field in the same model structure as an effect of percent concealing vegetation, because these covariates are both related to measures of the same potentially concealing vegetative cover.

#### Light Effect (l)

There was not enough variation in the light covariate to include it in any models for most of the data sets. I did include an additive parameter for light in some models in the candidate model set fit to the Black Mountain fall 2014 data; this was modeled as an additive parameter for those observations recorded in “flat” light.

#### Variance Estimation for Huggins Model Analyses

I modeled variance in the estimates of population size,  $Var(\bar{N})$ , using the delta method for variance approximation (Seber, 1982). The sources of variance in this estimator were variance in the estimated number of groups missed by all observers,  $\bar{G}_{missed}$ , and variance in the average estimated group size of groups missed,  $\bar{n}_{missed}$ .

$$Var(\bar{N}) = (\bar{G}_{missed} * Var(\bar{n}_{missed}))^2 * (\bar{n}_{missed} * Var(\bar{G}_{missed}))^2$$

I based  $Var(\bar{n}_{missed})$  for each data set on the estimated values of  $\bar{n}_{j,missed}$  from each of the observed groups in the data set. I used model-averaged estimates of  $Var(\bar{G}_{missed})$  from outputs in program MARK (White and Burnham, 1999). I found the 90-percent confidence intervals for the estimate of  $\bar{N}$  based on the estimated  $Var(\bar{N})$ , applying a log-normal distribution to obtain the lower and upper confidence limits. This approach has been used before to model uncertainty in Huggins method simultaneous double-observer analyses (Griffin and others, 2009).



## Comparisons of Estimated Abundance

It is not possible to make any comparisons of the estimated burro populations under any of the three types of analysis (L-P 1, L-P 2, and Huggins) to a “true” burro population size. In the future, it would be possible to test the efficacy of these and other analytical methods such as distance sampling (Buckland and others, 2001) or the use of infrared (IR; Beaver and others, 2014), if there are burro populations of known size, if there can be pre-gather and post-gather surveys (BLM, written commun., 2000a, 200b), or if there are burro populations with a substantial number of radio-collared animals (for example, Griffin and others, 2013).

## Results

### Areas Surveyed

Aerial surveys covered the entirety of Lake Pleasant HMA, McGee Mountain HMA, Warm Springs Canyon HMA, and Sinbad HMA. As much as possible, surveys were designed so that the spatial extent covered not only the HMA of interest, but also any surrounding lands where burros were known to possibly occur. Only portions of Cibola Trigo HMA, Black Mountain HMA, and High Rock HMA were covered.

The area surveyed in Cibola Trigo was within  $\approx 5$ –10 km (3–6 mi) of the Colorado River. An unknown number of burros were outside the surveyed area of Cibola-Trigo HMA at the time of the survey. That number cannot be estimated from these data without making explicit assumptions about the spatial distribution of burros in those areas, but the survey timing was chosen with the intent to maximize the proportion of the total population that was close to the Colorado River, based on past observations of seasonal burro movements.

The area surveyed at Black Mountain HMA in the spring was north of Arizona Highway 68 and south of U.S. Highway 93, whereas the fall surveys covered the portion of Black Mountain HMA south of Highway 68. Some areas of Black Mountain HMA were not surveyed, because of proximity to urban development or because high winds limited flight time.

The surveys of the burro population using Warm Springs Canyon HMA and High Rock HMA were designed to entirely survey Warm Springs Canyon HMA and that portion of High Rock HMA where it was thought that burros may also have occurred. The portion of the High Rock HMA surveyed was within  $\approx 4.5$  km ( $\approx 2.8$  mi) of the eastern boundary and included all areas contiguous with Warm Springs Canyon HMA where burros had been previously reported to, or observed by, BLM Surprise District and BLM Winnemucca District staff.

### Observed Burro Groups

Sample sizes of observations usable for L-P analysis estimation and Huggins model fitting varied (table 4). For Lake Pleasant HMA and Black Mountain HMA, the two Arizona HMAs where distance to each burro group was measured in 100-m increments, the sample size of double-observer observations within 200 m of the transect line was far smaller than the entire data set. Distance was also measured in 100-m increments at Warm Springs Canyon HMA and McGee Mountain HMA, but the surveys there were not designed to be exclusively parallel transects.

**Table 4.** Sample size of burro groups observed (observations) and simultaneous double-observer observations that were useful for fitting Huggins models. For Lake Pleasant and Black Mountain herd management areas only, the number of observations within 200 meters of a transect is also listed.

Study area	Observations	Double-observer observations	Double-observer observations $\leq 200$ meters
Lake Pleasant	50	46	26
Cibola-Trigo	130	128	n/a
Black Mountain, spring	68	66	36
Black Mountain, fall	229	228	157
Sinbad	23	23	n/a
Warm Springs Canyon and McGee Mountain <sup>1</sup>	29	25	n/a

<sup>1</sup>This data set includes observations from both herd management areas.

For all those data sets where distance was measured in 100-m increments, the most frequent distance category varied (fig. 1). In three of these four data sets, the most common observation distance was not the closest category ( $<100$  m). This result is consistent with the possibility that burros do move away from the sound of the approaching helicopter or also with the possibility that overall detection probability is lowered when burros are located immediately below the helicopter flight path. The latter phenomenon may arise if observers pay attention to the middle of the area searched and also because the back-seat observers have effectively no view immediately under the helicopter.

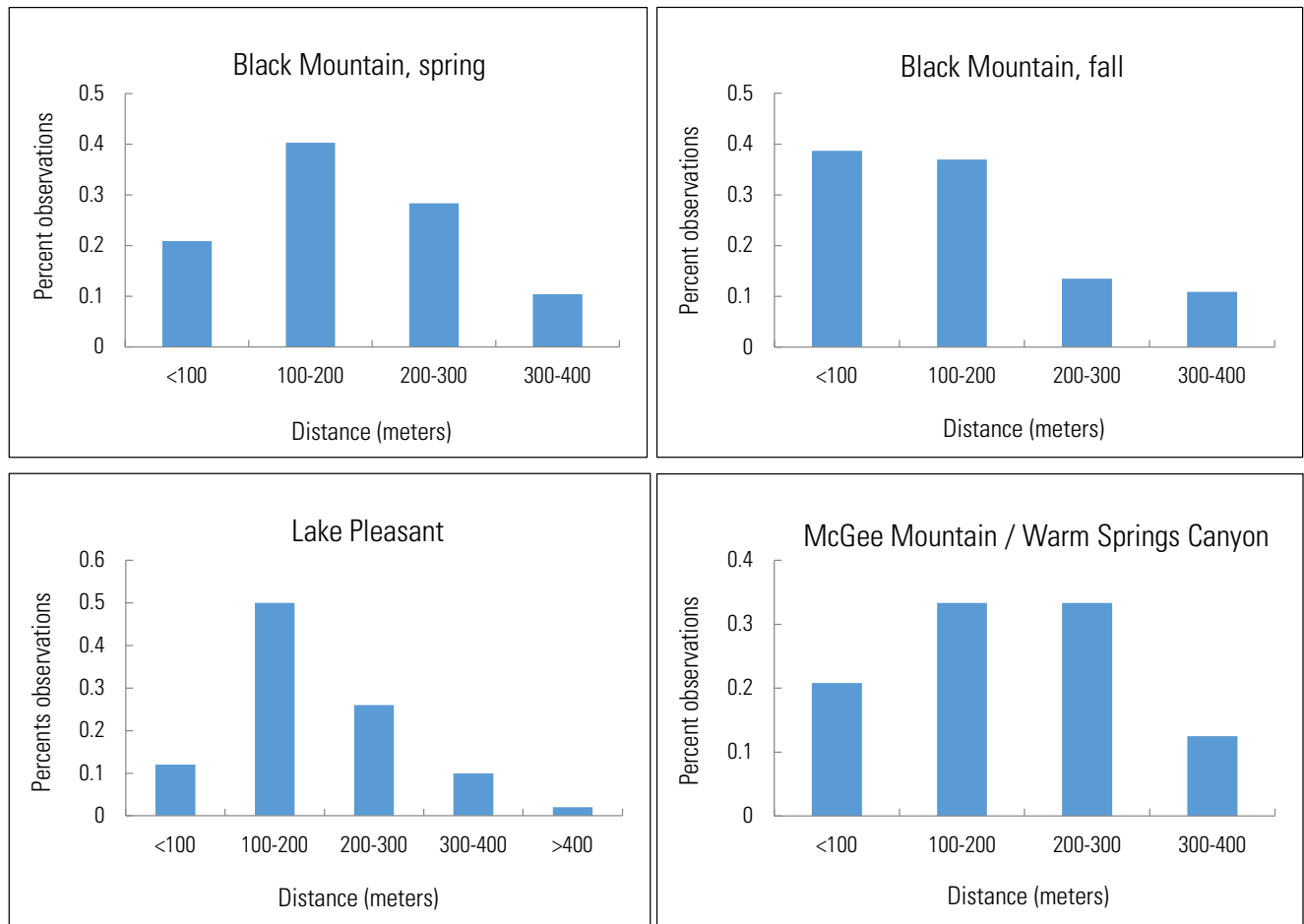
For those four data sets in which distance was recorded in 100-m increments, I found little to no correlation between the natural log of burro group size and distance (Black Mountain HMA spring,  $n = 68$ ,  $\rho = -0.163$ ; Black Mountain HMA fall,  $n = 229$ ,  $\rho = 0.054$ ; Lake Pleasant HMA,  $n = 50$ ,  $\rho = 0.002$ ; McGee Mountain HMA and Warm Springs Canyon HMA,  $n = 29$ ,  $\rho = 0.045$ ).

The distribution of observed burro group distance from the Colorado River in the Cibola-Trigo HMA surveys indicated that a fraction of the burro population was found at distances up to 10 km from the river (fig. 2). In some parts of the Trigo Mountains, the extent of the survey was only  $\approx 6$  km from the river.

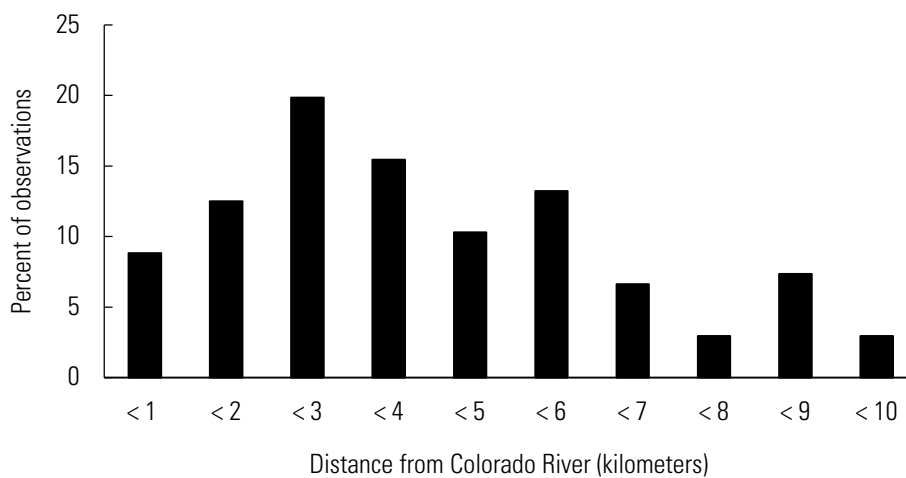
### Huggins Model Weights and Covariate Effects

Because analysis with the Huggins modeling approach fits models to observed data sets, the resulting model weights differed from data set to data set (table 5).

For the Black Mountain HMA spring survey data set and the McGee Mountain HMA and Warm Springs Canyon HMA data sets, the simplest model considered (M1) garnered the greatest weight of evidence. As with all models considered, this model was one with effects of group size and the burro group being on the pilot's side of the flight line. At Sinbad, the highest ranking model (M9) also had separate detection probability intercepts for front- and back-seat observers. At Lake Pleasant HMA, the highest ranking model (M17) had front- and back-seat observers sharing a detection probability intercept and also included a back-seat observer effect. At Cibola Trigo HMA, the highest ranking model (M33) had one shared intercept for front- and back-seat observers and also included effects of activity and visual field. In the Black Mountain HMA fall surveys, the highest ranked of the models considered had separate intercepts for front- and back-seat observers and also included effects of visual field and flat lighting. I caution against interpreting the top-ranked model to be the single "best" model, though. Small differences in AICc scores between many model structures meant that the model-averaged population estimates drew heavily from the estimates of many models.



**Figure 1.** Frequency distribution of recorded observation distances for the four data sets where distance was recorded in 100-meter (m) increments. The frequency in each distance category is expressed as a percentage of the total number of observations (percent observations). [>, greater than]



**Figure 2.** Distribution of distances from the Colorado River for burro group observations made in the Cibola-Trigo herd management area survey. Distances are approximate because they are based on location of the helicopter at time of observation. [<, less than]

**Table 5.** Model weights ( $w_k$ ) and differences in Akaike's Information Criterion scores ( $\Delta AICc$ ) for models fit to each data set. Higher model weights and lower  $\Delta AICc$  scores are indicative of greater support for that model, given the data and the models considered. For each data set, the model in bold is the one model with the greatest weight of evidence, given the data and the candidate model set considered. Model structure descriptions are listed in table 3.

Model number	Lake Pleasant		Cibola-Trigo		Black Mountain, spring		Black Mountain, fall*		Bi-State <sup>1</sup>		Sinbad	
	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$
M1	0.028	4.13	0.013	4.93	<b>0.076</b>	<b>0.00</b>	*	18.88	<b>0.389</b>	<b>0.00</b>	0.019	6.04
M2			0.010	5.34	0.033	1.64	*	20.18			0.006	8.37
M3					0.012	3.74	*	21.24				
M4					0.006	5.24	*	21.83				
M5			0.006	6.50	0.016	3.07	*	20.86			0.002	10.89
M6	0.018	5.05			0.027	2.05	*	20.23				
M7	0.010	6.24	0.008	5.81	0.040	1.27	*	19.21	0.145	1.98	0.006	8.45
M8	0.006	7.28	0.005	6.62	0.014	3.39	*	20.40				
M9	0.018	5.09			0.030	1.84	0.012	3.44	0.121	2.34	<b>0.380</b>	<b>0.00</b>
M10			0.004	7.06	0.013	3.51	0.005	5.43			0.119	2.32
M11					0.005	5.64	0.008	4.30				
M12					0.002	7.19	0.005	5.34				
M13			0.002	8.27	0.006	5.00	0.005	5.23			0.034	4.84
M14	0.011	6.02			0.011	3.92	0.008	4.31				
M15	0.006	7.20	0.003	7.56	0.016	3.16	0.006	4.83	0.042	4.44	0.114	2.40
M16	0.003	8.31			0.005	5.32	0.016	2.92				
M17	0.223	<b>0.00</b>	0.002	8.52	0.040	1.28	0.009	4.11	0.173	1.62	0.023	5.62
M18			0.002	8.91	0.017	2.95	0.006	4.73			0.007	8.14
M19					0.006	5.03	0.004	5.64				
M20					0.003	6.41	0.006	4.89				
M21			0.001	10.11	0.009	4.24	0.003	6.00			0.002	10.67
M22	0.149	0.80			0.014	3.31	*	8.11				
M23	0.099	1.62	0.001	9.45	0.023	2.39	0.004	5.61	0.060	3.75	0.007	8.04
M24	0.050	3.01			0.008	4.52	*	8.74				
M25	0.077	2.12	0.001	10.17	0.014	3.44	0.006	4.94	0.053	4.00	0.171	1.60
M26			0.001	10.56	0.006	5.14	*	7.36			0.049	4.10
M27					0.002	7.26	0.004	5.92				
M28					0.001	8.67	*	7.35				
M29			0.000	11.85	0.003	6.46	0.008	4.24			0.013	6.70
M30	0.050	2.98			0.005	5.50	0.005	5.24				
M31	0.034	3.78	0.001	11.18	0.008	4.58	*	7.51	0.017	6.22	0.049	4.08
M32	0.016	5.23			0.003	6.74	0.025	2.03				
M33			<b>0.148</b>	<b>0.00</b>	0.027	2.06	0.016	2.87				
M34			0.070	1.48	0.014	3.43	0.008	4.20				

**Table 5.** Model weights ( $w_k$ ) and differences in Akaike's Information Criterion scores ( $\Delta AICc$ ) for models fit to each data set. Higher model weights and lower  $\Delta AICc$  scores are indicative of greater support for that model, given the data and the models considered. For each data set, the model in bold is the one model with the greatest weight of evidence, given the data and the candidate model set considered. Model structure descriptions are listed in table 3.—Continued

Model number	Lake Pleasant		Cibola-Trigo		Black Mountain, spring		Black Mountain, fall*		Bi-State <sup>1</sup>		Sinbad	
	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$
M35	0.012	5.82	0.062	1.72	0.043	1.14	0.013	3.29				
M36	0.004	8.06	0.036	2.82	0.024	2.27	0.008	4.39				
M37			0.061	1.76	0.010	3.96	0.014	3.16				
M38			0.029	3.27	0.005	5.38	0.011	3.67				
M39	0.007	6.79	0.026	3.46	0.017	3.02	0.028	1.81				
M40	0.002	9.08	0.015	4.60	0.009	4.21	*	22.10				
M41			0.025	3.53	0.016	3.16	*	22.71				
M42			0.012	5.06	0.009	4.35	*	20.84				
M43	0.097	1.66	0.010	5.32	0.024	2.27	*	21.09				
M44	0.035	3.71	0.006	6.46	0.015	3.21	0.003	6.25				
M45			0.011	5.20	0.005	5.38	0.003	6.18				
M46			0.005	6.79	0.003	6.55	0.006	4.81				
M47	0.033	3.83	0.004	6.99	0.008	4.46	0.002	6.65				
M48	0.011	5.93	0.002	8.20	0.005	5.38	*	8.66				
M49			0.006	6.37	0.036	1.50	0.002	6.75				
M50			0.004	7.06	0.016	3.15	*	9.31				
M51					0.005	5.31	0.003	6.05				
M52					0.003	6.76	*	7.67				
M53			0.003	8.14	0.008	4.55	0.005	5.41				
M54					0.012	3.60	*	7.90				
M55			0.004	7.17	0.019	2.72	0.005	5.12				
M56					0.007	4.90	0.009	4.07				
M57			0.003	8.06	0.010	4.00	0.010	3.79				
M58			0.002	8.77	0.006	5.06	*	20.66				
M59					0.002	7.24	*	22.05				
M60					0.001	8.76	*	23.22				
M61			0.001	9.91	0.003	6.51	*	23.79				
M62			0.002	8.92	0.005	5.51	*	22.70				
M63					0.007	4.66	*	22.11				
M64					0.002	6.87	*	20.98				
M65			0.001	10.01	0.019	2.74	*	22.28				
M66			0.001	10.67	0.008	4.42	0.003	6.23				
M67					0.003	6.57	*	7.36				
M68					0.001	7.89	0.002	7.24				

**Table 5.** Model weights ( $w_k$ ) and differences in Akaike's Information Criterion scores ( $\Delta AICc$ ) for models fit to each data set. Higher model weights and lower  $\Delta AICc$  scores are indicative of greater support for that model, given the data and the models considered. For each data set, the model in bold is the one model with the greatest weight of evidence, given the data and the candidate model set considered. Model structure descriptions are listed in table 3.—Continued

Model number	Lake Pleasant		Cibola-Trigo		Black Mountain, spring		Black Mountain, fall*		Bi-State <sup>1</sup>		Sinbad	
	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$
M69			0.000	11.80	0.004	5.66	0.003	6.03				
M70			0.001	10.85	0.007	4.83	0.004	5.89				
M71					0.011	3.78	0.002	6.96				
M72					0.004	5.97	0.003	6.21				
M73			0.000	11.70	0.005	5.64	*	8.14				
M74			0.000	12.36	0.003	6.65	*	9.76				
M75					0.001	8.83	*	10.35				
M76					0.000	10.16	0.003	6.20				
M77			0.000	13.57	0.001	7.90	*	8.79				
M78					0.002	7.06	*	8.79				
M79			0.000	12.62	0.004	5.98	0.005	5.26				
M80					0.001	8.20	0.002	6.91				
M81			0.100	0.78	0.009	4.21	*	8.94				
M82			0.050	2.16	0.005	5.59	0.011	3.57				
M83			0.052	2.08	0.016	3.16	0.007	4.56				
M84			0.032	3.06	0.009	4.30	0.006	4.81				
M85			0.042	2.53	0.004	6.15	0.005	5.33				
M86			0.021	3.94	0.002	7.59	0.013	3.32				
M87			0.022	3.81	0.006	5.07	*	24.09				
M88			0.013	4.82	0.003	6.28	*	24.71				
M89			0.017	4.37	0.005	5.34	*	22.71				
M90			0.008	5.79	0.003	6.55	*	23.01				
M91			0.008	5.73	0.003	6.54	*	8.27				
M92			0.005	6.75	0.005	5.27	*	8.21				
M93			0.007	6.09	0.002	7.60	*	7.33				
M94			0.003	7.57	0.001	8.79	0.002	6.76				
M95			0.004	7.46	0.003	6.54	*	10.20				
M96			0.002	8.54	0.002	7.47	*	10.88				
M97							*	8.94				
M98							0.002	7.21				
M99							*	9.26				
M100							0.002	6.65				
M101							0.004	5.59				
M102							0.024	2.08				
M103							0.012	3.48				

**Table 5.** Model weights ( $w_k$ ) and differences in Akaike's Information Criterion scores ( $\Delta AICc$ ) for models fit to each data set. Higher model weights and lower  $\Delta AICc$  scores are indicative of greater support for that model, given the data and the models considered. For each data set, the model in bold is the one model with the greatest weight of evidence, given the data and the candidate model set considered. Model structure descriptions are listed in table 3.—Continued

Model number	Lake Pleasant		Cibola-Trigo		Black Mountain, spring		Black Mountain, fall*		Bi-State <sup>1</sup>		Sinbad	
	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$
M104							0.007	4.68				
M105							0.014	3.15				
M106							0.012	3.54				
M107							0.035	1.33				
M108							0.019	2.60				
M109							0.053	0.51				
M110							0.026	1.90				
M111							0.013	3.36				
M112							0.014	3.17				
M113							0.028	1.80				
M114							0.023	2.20				
M115							<b>0.068</b>	<b>0.00</b>				
M116							0.032	1.53				
M117							0.004	5.50				
M118							0.005	5.10				
M119							0.009	4.12				
M120							0.013	3.30				
M121							0.009	3.97				
M122							0.010	3.82				
M123							0.019	2.59				
M124							0.025	2.03				
M125							0.009	4.14				
M126							0.004	5.56				
M127							0.002	6.75				
M128							0.003	6.32				
M129							0.005	5.23				
M130							0.012	3.39				
M131							0.007	4.68				
M132							0.020	2.47				
M133							0.010	3.90				
M134							0.005	5.43				
M135							0.010	3.80				
M136							0.008	4.21				
M137							0.026	1.96				
M138							0.012	3.56				
M139							*	7.57				

**Table 5.** Model weights ( $w_k$ ) and differences in Akaike’s Information Criterion scores ( $\Delta AICc$ ) for models fit to each data set. Higher model weights and lower  $\Delta AICc$  scores are indicative of greater support for that model, given the data and the models considered. For each data set, the model in bold is the one model with the greatest weight of evidence, given the data and the candidate model set considered. Model structure descriptions are listed in table 3.—Continued

Model number	Lake Pleasant		Cibola-Trigo		Black Mountain, spring		Black Mountain, fall*		Bi-State <sup>1</sup>		Sinbad	
	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$	$w_k$	$\Delta AICc$
M140							0.002	7.17				
M141							0.003	6.19				
M142							0.005	5.37				
M143							0.007	4.54				
M144							0.009	4.05				

\*Only the top 96-ranked models from the Black Mountain fall surveys were used in model averaging. Models indicated with an asterisk were ranked lower than 96th, so no model weight is listed for them.

<sup>1</sup>The Bi-State surveys included observations from McGee Mountain herd management area (HMA) and Warm Springs Canyon HMA.

Meaningful measures of any given sighting covariate’s effect on detection probability can be seen after model averaging, in comparison to a baseline detection probability (table 6). In all data sets, the effect of a burro group being on the pilot’s side of the survey path was negative; such groups were estimated to have lower detection probabilities for the front-seat-observer pair ( $\hat{p}_{F,j}$ ) and correspondingly higher correction factors ( $\bar{\theta}_j$ ). The magnitude of the pilot’s-side effect varied from pilot to pilot. Back-seat-observer acuity also varied. The lowest detection probabilities (and highest correction factors) were for groups on the pilot’s side of the flight line when there was a back-seat observer with relatively low acuity.

**Table 6.** Estimated baseline detection probabilities and effects of modeled sighting covariates. Model-averaged estimated detection probabilities are shown for front-seat observers ( $\hat{p}_{F,j}$ ), back-seat observers ( $\hat{p}_{B,j}$ ), and overall detection probability across all observers ( $\hat{p}_{\bullet,j}$ ). The model-averaged group-specific correction factor for a given set of covariates is shown as  $\bar{\theta}_j$ . For each data set considered, only those effects are shown that were included in model structures considered. The baseline detection probability shown, in bold, is always for a group size of four animals, not moving, not on the pilot’s side, at a distance of 100–200 meters (or  $<1/4$  mile [ $<400$  meters] where distance was recorded in  $1/4$ -mile increments), and under high-contrast light.

[RB, AGFD 1, CB, JH, CW, SS, PG, and DC are initials of individual observers noted in table 2.]

Data set	Effect description	$\hat{p}_{F,j}$	$\hat{p}_{B,j}$	$\hat{p}_{\bullet,j}$	$\bar{\theta}_j$
Lake Pleasant					
	Baseline (back observer RB)	<b>0.52</b>	<b>0.53</b>	0.77	1.32
	Pilot Side <sup>1</sup>	0.05	<b>0.53</b>	0.55	1.87
	Back Observer AGFD 1	<b>0.52</b>	0.20	0.62	1.65
	Pilot Side <sup>1</sup> and Back Observer AGFD 1	0.05	0.20	0.24	4.5
	Effect of visual field being not “open”	0.55	0.55	0.80	1.29



**Table 6.** Estimated baseline detection probabilities and effects of modeled sighting covariates. Model-averaged estimated detection probabilities are shown for front-seat observers ( $\hat{p}_{F,j}$ ), back-seat observers ( $\hat{p}_{B,j}$ ), and overall detection probability across all observers ( $\hat{p}_{\bullet,j}$ ). The model-averaged group-specific correction factor for a given set of covariates is shown as  $\bar{\theta}_j$ . For each data set considered, only those effects are shown that were included in model structures considered. The baseline detection probability shown, in bold, is always for a group size of four animals, not moving, not on the pilot’s side, at a distance of 100–200 meters (or <¼ mile [ $<400$  meters] where distance was recorded in ¼-mile increments), and under high-contrast light.—Continued  
 [RB, AGFD 1, CB, JH, CW, SS, PG, and DC are initials of individual observers noted in table 2.]

Data set	Effect description	$\hat{p}_{F,j}$	$\hat{p}_{B,j}$	$\hat{p}_{\bullet,j}$	$\bar{\theta}_j$
Black Mountain, spring 2014					
	Baseline (back observer anyone but CB)	<b>0.72</b>	<b>0.73</b>	0.92	1.08
	Pilot Side	0.41	<b>0.73</b>	0.84	1.19
	Back Observer CB	<b>0.72</b>	0.76	0.93	1.07
	Pilot Side and Back Observer CB	0.41	0.76	0.86	1.17
	Effect of visual field being not “open”	0.67	0.68	0.90	1.12
	Effect of animals moving	0.76	0.76	0.94	1.07
Black Mountain, fall 2014					
	Baseline (back observer CB)	<b>0.85</b>	<b>0.69</b>	0.95	1.05
	Pilot Side	0.70	<b>0.69</b>	0.91	1.11
	Back Observer AGFD 7	<b>0.85</b>	0.60	0.94	1.06
	Back Observer AGFD 8	<b>0.85</b>	0.58	0.94	1.06
	Pilot Side and Back Observer AGFD 7	0.70	0.60	0.88	1.11
	Pilot Side and Back Observer AGFD 8	0.70	0.58	0.88	1.15
	Effect of visual field being not “open”	0.85	0.69	0.95	1.05
	Effect of animals moving	0.86	0.71	0.96	1.04
	Effect of flat light	0.79	0.60	0.91	1.09
Cibola-Trigo					
	Baseline (back observer JH)	<b>0.82</b>	<b>0.81</b>	0.96	1.04
	Pilot Side	0.58	<b>0.81</b>	0.92	1.09
	Back Observer AGFD 9	<b>0.82</b>	0.81	0.97	1.04
	Back Observer CW	<b>0.82</b>	0.82	0.97	1.04
	Pilot Side and Back Observer AGFD 9	0.58	0.81	0.92	1.09
	Pilot Side and Back Observer CW	0.58	0.82	0.92	1.09
	Effect of visual field being not “Open”	0.62	0.61	0.85	1.18
	Effect of animals moving	0.88	0.88	0.99	1.02
McGee Mountain and Warm Springs Canyon					
	Baseline (back observer SS)	<b>0.72</b>	<b>0.72</b>	0.92	1.09
	Pilot Side	0.47	<b>0.72</b>	0.85	1.18
	Back Observer PG	<b>0.72</b>	0.77	0.93	1.07
	Pilot Side and Back Observer PG	0.47	0.77	0.88	1.14

**Table 6.** Estimated baseline detection probabilities and effects of modeled sighting covariates. Model-averaged estimated detection probabilities are shown for front-seat observers ( $\hat{p}_{F,j}$ ), back-seat observers ( $\hat{p}_{B,j}$ ), and overall detection probability across all observers ( $\hat{p}_{\bullet,j}$ ). The model-averaged group-specific correction factor for a given set of covariates is shown as  $\bar{\theta}_j$ . For each data set considered, only those effects are shown that were included in model structures considered. The baseline detection probability shown, in bold, is always for a group size of four animals, not moving, not on the pilot's side, at a distance of 100–200 meters (or <1/4 mile [<400 meters] where distance was recorded in 1/4-mile increments), and under high-contrast light.—Continued  
 [RB, AGFD 1, CB, JH, CW, SS, PG, and DC are initials of individual observers noted in table 2.]

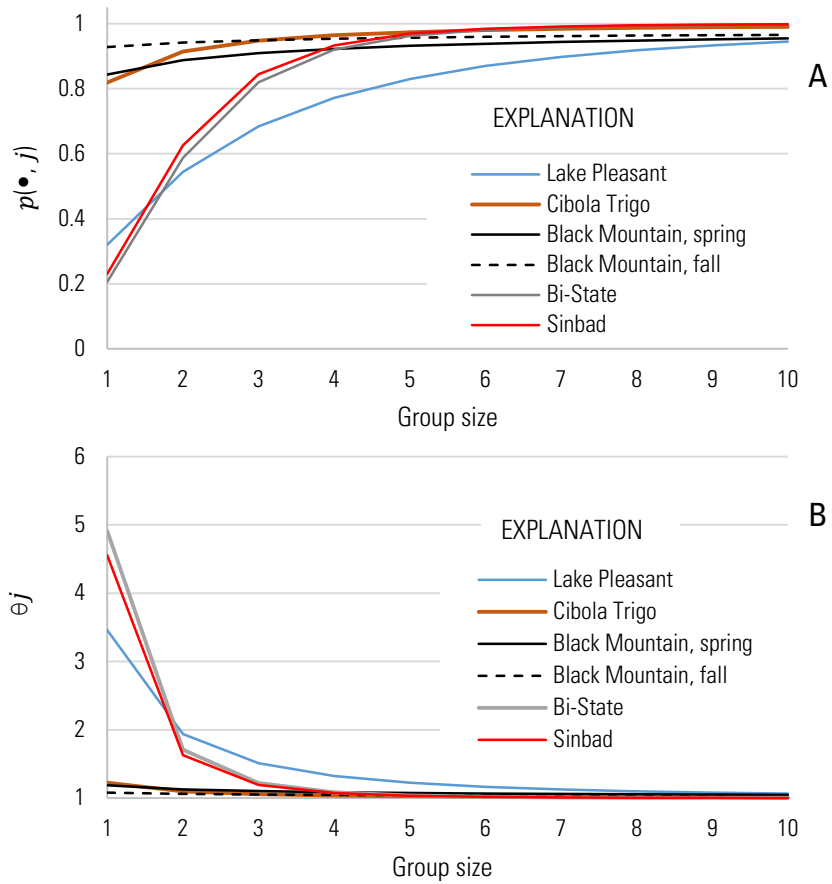
Data set	Effect description	$\hat{p}_{F,j}$	$\hat{p}_{B,j}$	$\hat{p}_{\bullet,j}$	$\bar{\theta}_j$
Sinbad					
	Baseline (back observer DC)	<b>0.74</b>	<b>0.74</b>	0.93	1.07
	Pilot Side	0.51	<b>0.74</b>	0.87	1.15
	Back Observer PG	<b>0.74</b>	0.8	0.95	1.06
	Pilot Side and Back Observer PG	0.51	0.8	0.9	1.12
	Effect of animals moving	0.69	0.69	0.9	1.11

<sup>1</sup>The pilot was encouraged to not observe burros on the Lake Pleasant survey.

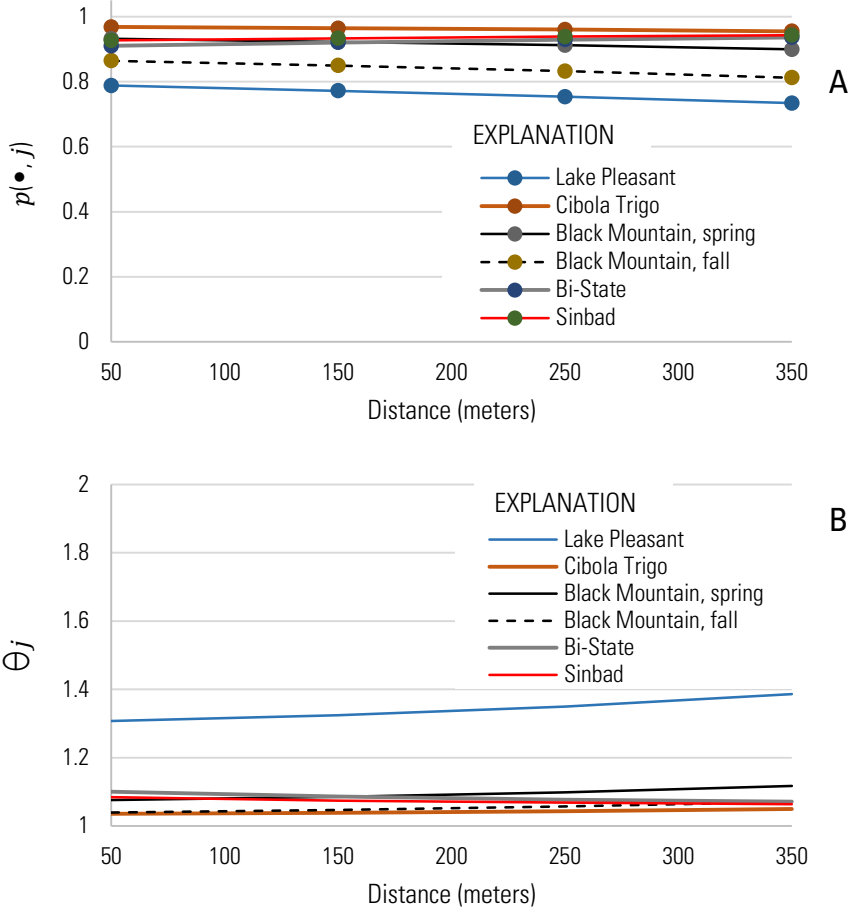
In all data sets, detection probability was predicted to increase as a function of group size (fig. 3). Correspondingly, the group-specific correction factors applied to each group decreased as group size increased.

For those data sets where the effect of distance was included as a potential covariate, burro groups at greater distances had only slightly lower estimated detection probabilities (fig. 4). Similarly, for data sets in which the effect of vegetative cover or the effect of “broken” or “tree” visual fields was considered, only very slightly lower detection probability was predicted for burro groups in increasing vegetative cover at Black Mountain HMA (fig. 5), and lower detection probability was predicted where the visual field was anything but “open” (table 6). The effect of vegetative cover was the opposite direction in the Lake Pleasant HMA data set (fig. 5); this may have been because there was a slight, but not statistically significant, negative correlation between vegetative cover and distance in that data set ( $r = -0.216$ ,  $P = 0.133$ , degrees of freedom = 48); the observations in higher vegetative cover tended to be made at closer distances.

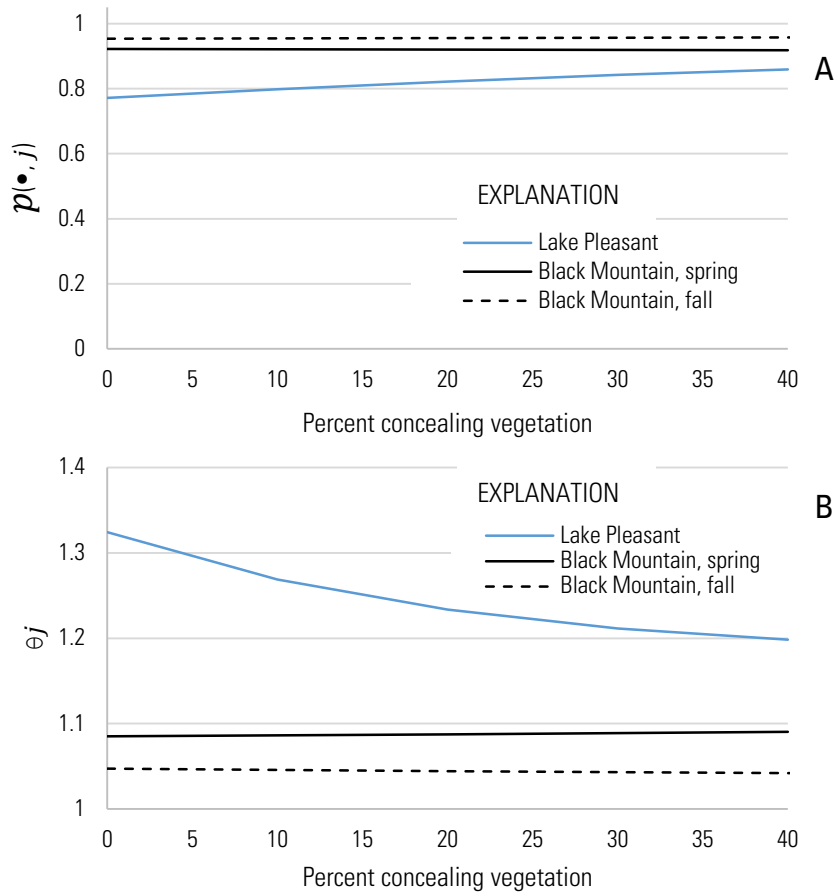
For two data sets from the Black Mountain HMA, there were some differences in estimated detection probabilities. It is not possible to say whether these differences were due to the high winds in the spring surveys, due to different sighting conditions in the two seasons, or due to differences in pilots and observers.



**Figure 3.** Expected effect of group size on (A) group-specific estimated overall detection probability ( $\hat{p}_{\bullet, j}$ ) and (B) correction factor ( $\bar{\theta}_j$ ) in six simultaneous double-observer data sets from 2014 burro aerial surveys. The expected values are for stationary burro groups in the open, 100–200 meters from the transect line, not on the pilot’s side of the flight line, and with the intercept back-seat observer.



**Figure 4.** Expected effect of distance on (A) group-specific estimated overall detection probability ( $\hat{p}_{\bullet, j}$ ) and (B) correction factor ( $\bar{\theta}_j$ ) in six simultaneous double-observer data sets from 2014 burro aerial surveys. The expected values are for stationary burro groups in the open, with group size of four, not on the pilot's side of the flight line, with no concealing vegetation, and with the intercept back-seat observer.



**Figure 5.** Expected effect of the percentage of concealing vegetation on (A) group-specific estimated overall detection probability ( $\hat{p}_{\bullet, j}$ ) and (B) correction factor ( $\bar{\theta}_j$ , lower graph) in three simultaneous double-observer data sets from 2014 burro aerial surveys. Values shown are for stationary burro groups in the open, with group size of four, not on the pilot's side of the flight line, at distances of 100–200 meters, and with the intercept back-seat observer. The vegetation effect was not modeled for the Cibola-Trigo herd management area (HMA), Sinbad HMA, or McGee Mountain HMA and Warm Springs Canyon HMA data sets.

### Estimated Abundance

Estimated population abundance values for each HMA were not entirely consistent across the three different estimators making use of all available observation data, nor in comparing those to estimates made with reference only to observations within 200 m of the transect (table 7). Without known population sizes of burros, I cannot make conclusions about which estimator provides the most accurate estimates of population size.

I note that the L-P 1 point estimate for burro population size at Sinbad HMA was *lower* than the number of burros seen. This result is largely because the estimator uses an arithmetic average group size ( $\bar{n}_{observed}$ ) in calculations, whereas the fairly small number of observed groups at Sinbad ( $n=23$ ) included several large groups of 12, 13, 27, and 30 burros.

Inferences from all estimates in table 7 are limited to the time of survey and the spatial extent of the survey.

**Table 7.** Observed numbers of burros (minimum number of animals [MNA]) and estimated abundance of surveyed burro populations (in bold font). Estimates ( $\hat{N}$ ) are from three types of analysis: Lincoln-Petersen technique 1 (L-P 1), Lincoln-Petersen technique 2 (L-P 2), and Huggins analysis using sighting covariates (Huggins). Where possible,  $\hat{N}$  for L-P 1 and L-P 2 is also shown based only on observations  $\leq 200$  meters from the transect line, with an area correction factor of two. Ninety percent confidence intervals are in braces. Estimated percentages of burros present but not seen during the survey (Est. % Missed) and ratios of foals to adults are based on the Huggins analysis.

[HMA, herd management area; <, less than]

Survey area	MNA	All observations; no area correction		Observations <200 meters; area correction = 2		Huggins $\hat{N}^3$	Estimated percent missed	Estimated foals per adult <sup>4</sup>
		L-P 11 $\hat{N}$	L-P 22 $\hat{N}$	L-P 11 $\hat{N}^5$	L-P 22 $\hat{N}^5$			
Lake Pleasant	228	<b>350</b>	<b>306</b>	<b>428</b> <sup>5</sup>	<b>257</b> <sup>5</sup>	<b>460</b>	50.4	0.11
HMA		{240 <sup>7</sup> , 647}	{245 <sup>7</sup> , 408}	{268, 1067}	{212, 314}	{281, 1,239}		
Black	258	<b>293</b>	<b>245</b>	<b>325</b> <sup>5</sup>	<b>305</b> <sup>5</sup>	<b>296</b>	12.7	0.17
Mountain		{234 <sup>7</sup> , 392}	{220 <sup>7</sup> , 276}	{255 <sup>7</sup> , 448}	{279, 336}	{274, 344}		
HMA, spring <sup>6</sup>								
Black	1,109	<b>1,453</b>	<b>1,582</b>	<b>2,145</b> <sup>5</sup>	<b>2,361</b> <sup>5</sup>	<b>1,221</b>	9.1	0.21
Mountain		{1,260,	{1,511,	{1,842,	{2,250,	{1,170,		
HMA, fall <sup>6</sup>		1,716}	1,659}	2,567}	2,484}	1,313}		
Cibola-Trigo	365	<b>427</b>	<b>487</b>	—	—	<b>457</b>	20.2	0.04
HMA <sup>6</sup>		{362 <sup>7</sup> , 519}	{453 <sup>7</sup> , 528}			{410, 555}		
High Rock	0	<b>0</b>	<b>0</b>	—	—	<b>0</b>		—
HMA <sup>8</sup>								
Warm Springs	20	<b>24</b>	<b>42</b>	—	—	<b>34</b>	47.6	0.05
Canyon		{16 <sup>7</sup> , 49}	{31, 68}			{22, 113}		
HMA								
McGee	97	<b>168</b>	<b>146</b>	—	—	<b>124</b>	21.6	0.15
Mountain		{112, 341}	{117, 194}			{104, 202}		
HMA								
Sinbad HMA	160	<b>154</b>	<b>219</b>	—	—	<b>187</b>	14.7	0.06
		{114 <sup>7</sup> , 237}	{204, 237}			{185, 190}		

<sup>1</sup>Lincoln-Petersen technique L-P 1 applies a correction to groups seen by left-back (LB) and right-back (RB) observers based on the L-P estimated detection probability for LB.

<sup>2</sup>Lincoln-Petersen technique L-P 2 applies a correction to groups seen by left-front (LF) and LB observers based on the overall L-P estimated detection probability for both left observers.

<sup>3</sup>Huggins analysis method uses records of covariates from each observed group of burros.

<sup>4</sup>Based on Huggins analysis. Estimated ratios of foals to adults may not be indicative of true ratios of foals to adults given the seasonal timing of these surveys.

<sup>5</sup>Estimates based only on observations  $\leq 200$  meters from the transect. Half the survey area was included, so the area correction factor for these L-P estimators was 2.

<sup>6</sup>The survey design did not encompass the whole herd management area (HMA). Estimated abundance applies to the area surveyed at the time of the survey.

<sup>7</sup>The lower limit of the 90-percent confidence interval was lower than the number of animals seen.

<sup>8</sup>No burros were observed on the High Rock HMA.

## **Estimated Group Sizes of Missed Groups**

Huggins method population estimates in table 7 come from the sum of animals actually recorded and the estimated number of animals that were present, but missed. At Lake Pleasant HMA, McGee Mountain HMA and Warm Springs Canyon HMA, and Sinbad HMA, there were relatively high numbers of groups that were estimated to have been missed by all observers (fig. 6), but most of these were of small group size.

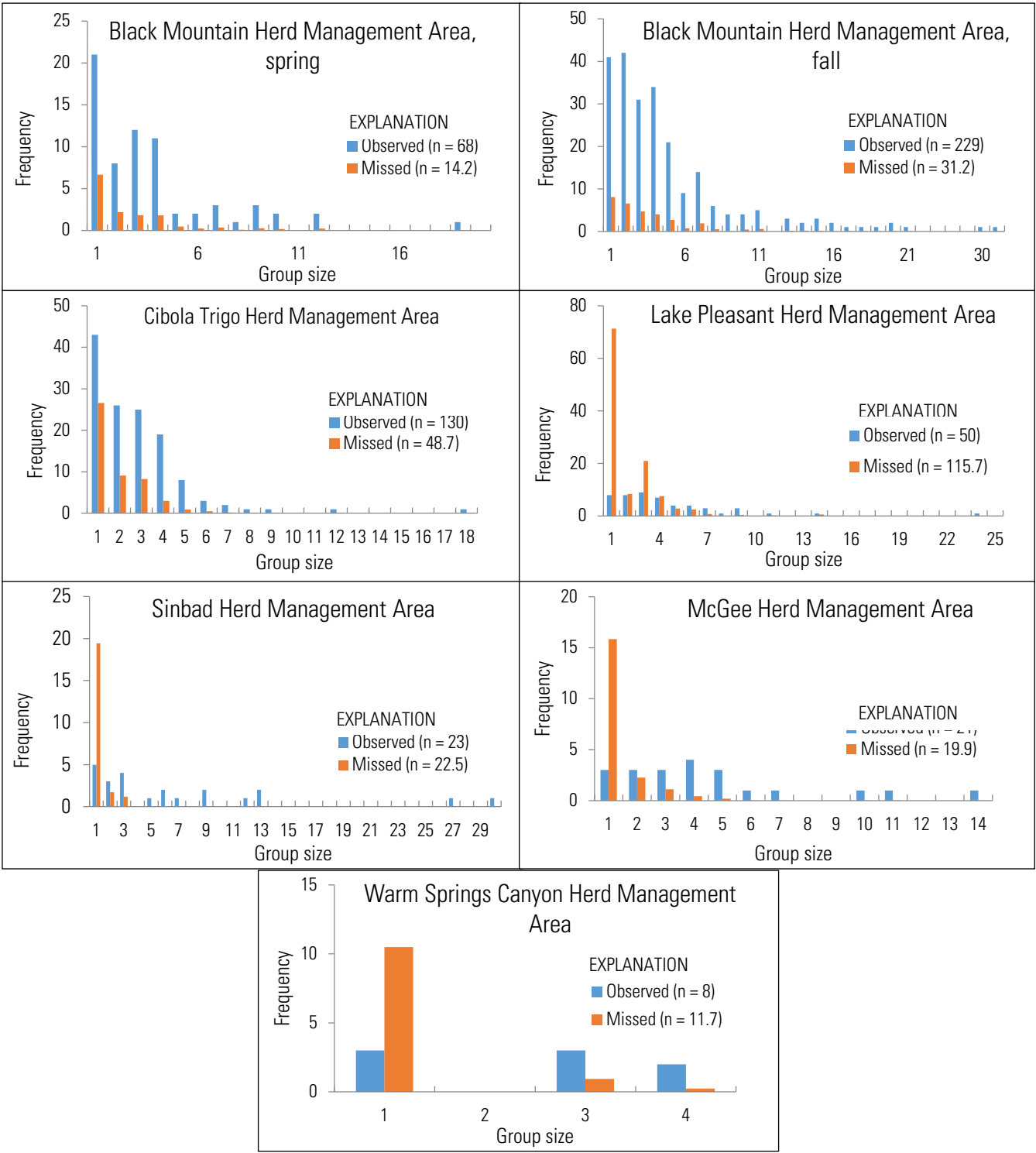
In Huggins analyses, correction factors applied to observed groups always increase as overall detection probability decreases. As a result of the positive correlation between group size and detection probability, the highest numbers of estimated, missed groups were always associated with observed groups of single animals. Such groups had even higher correction factors, if the observation was on the pilot's side of the flight line where detection probability was universally estimated to have been lower than if the group was on the left side of the flight line. To illustrate, of the 232 burros that were estimated to have been missed in the Lake Pleasant survey, 67 were estimated based on just six observed groups of one burro each, on the pilot's side of the transect line. The extremely low detection probability for the pilot on the Lake Pleasant survey is not surprising, because the pilot in that survey was instructed to not make burro observations.

## **Discussion**

Our results, using two general approaches to the mark-recapture analysis of simultaneous double-observer data in burro aerial surveys, indicate that there can be substantial variation in estimated abundance values for burros depending on the analysis method used (table 7). Because I do not know the true number of burros in these surveyed populations I am, unfortunately, not able to make any conclusions about which analytical method was the most accurate. I can refer to theory, though, to make speculative suggestions about these analyses based on some observations about patterns in the recorded data.

Unmodeled heterogeneity in simultaneous double-observer data will tend to cause estimates of detection probability to be biased high, such that the resulting abundance estimates will tend to be lower than the true number of animals present in the surveyed area (Pollock and Kendall, 1987; Walter and Hone, 2003; Southwell and others, 2007; Barker, 2008; Laake and others, 2011; Griffin and others, 2013). In theory, accounting for the effects of sighting covariates, as is done with Huggins type analyses, should reduce the amount of bias due to unmodeled heterogeneity and should, therefore, reduce the bias in estimated abundance compared to the simpler Lincoln-Petersen techniques (L-P 1 and L-P 2).

In our results, the Huggins estimates of population size were higher than the L-P 1 estimates based on all available observations at Lake Pleasant HMA, Black Mountain HMA spring surveys, Cibola Trigo HMA, Warm Springs Canyon HMA, and Sinbad HMA. Huggins estimates were higher than L-P 2 estimates at Lake Pleasant HMA and Black Mountain HMA spring surveys. At McGee Mountain HMA and Black Mountain HMA fall surveys, the Huggins estimates of population size were slightly lower than the L-P estimates.



**Figure 6.** Histograms of the numbers (n) of observed burro groups and estimated numbers of missed groups in each of six herd management areas (HMAs) surveyed for burros in 2014. Estimated numbers of missed groups are based on Huggins method analysis.



Differences in the L-P and Huggins estimates can be explained by the estimators used in those analyses. To estimate abundance, L-P 1 and L-P 2 estimators divide a number of observed groups by one observer's estimated detection probability, or a joint estimated detection probability of two observers, to find an estimated number of total groups, then multiply that number of total groups by the arithmetic average of all observed group sizes. The analyses that only use observations closer than 200 m also apply an area correction factor. There is no guarantee, though, that the group sizes of observed groups match the group size of burro groups that were, in fact, missed by all observers. One might expect that smaller groups would be missed at a greater rate than large groups, as supported by the Huggins method analysis (figs. 2 and 5). If this happens, then the group sizes of observed groups would tend to be larger, on average, than the group sizes of missed groups.

In contrast to the L-P estimators, the abundance estimate from the Huggins estimator comes from adding an estimated number of missed groups to the observed groups in a way that accounts for the specific group sizes of each group that was estimated to have been missed. The estimated numbers of missed groups (fig. 5) are derived from an analysis that accounts for the effect of covariates on each observed group's likely detection probability. Thus, large burro groups, such as some of those at Black Mountain HMA, McGee Mountain HMA, and Sinbad HMA, generally have high estimated detection probabilities in the Huggins analyses, so they represent very few additional burros in the Huggins analysis because the group-specific correction factors applied to each group are the inverse of the group's estimated detection probability.

At the other end of the group size spectrum, burro groups with very low estimated sighting probabilities can have high correction factors and larger effects on Huggins analysis confidence intervals. In the example of the Lake Pleasant HMA data set, five groups, each of a single burro, were observed on the pilot's side of the helicopter flight path, with a visual field of "tree" or "broken," in 30 percent concealing vegetation. Each of these groups had an overall estimated detection probability of less than 10 percent, with corresponding correction factors of over 10. This low overall detection probability can be partly explained by the fact that the pilot was encouraged to not observe burros on the Lake Pleasant survey. The same calculations, albeit with different values, explain why there were estimated to be high numbers of unseen, small groups at McGee Mountain HMA, Warm Springs Canyon HMA, and Sinbad HMA (fig. 6). In all these cases, there were some observations of group size of one, with the highest correction factors resulting when the observation was on the pilot's side.

Pilots are tasked with flying the aircraft safely and that must be their priority. Estimates of the pilot effect (table 6) do, however, indicate that there is a great deal of variation in the pilots' acuity in detecting burros while flying. In the Lake Pleasant HMA surveys, the front-seat observers' detection probability was estimated to be 47 percent lower when the burro group was on the pilot's side of the flight line for the hypothetical burro group examples shown in table 6. The very wide confidence intervals for the Huggins analysis of the Lake Pleasant HMA data set are at least partially due to those observations with very low estimated detection probabilities. This is not a result of a flaw in the Huggins analysis; rather, it is a reflection of low estimated detection probabilities for at least some of the observed burro groups in that data set.

The distribution of distances from the Colorado River for burro observations in the Cibola-Trigo HMA calls into question the adequacy of the survey design to lead to a population estimate for the entirety of the population using that HMA. The survey was designed to take advantage of burros being concentrated near the river during the time of the survey, but a non-

negligible fraction of the burro groups were observed even as far as 10 km from the river, which was further from the river than the extent of surveys in some areas of the Trigo Mountains. I would caution, therefore, that the estimated population size based on analysis of the Cibola-Trigo survey data may be an underestimate of the total population using the HMA, to an extent that is unquantifiable based on these data alone.

### **Observer Acuity**

It is always best to use observers that have high acuity in aerial surveys. Results in this report point to the importance of using observers with the greatest possible skill in observing animals and placing observers with high acuity directly behind the pilot, when possible. When a burro group is on the pilot's side of the flight line, the expected overall detection probability for that group ( $\hat{p}_{\cdot,j}$ ) will be increased if there is an observer with high acuity seated directly behind the pilot.

For simultaneous double-observer surveys and analyses, it is also desirable to limit the number of observers that contribute to a given data set. It should be a goal to have an adequate sample size of observations to parameterize each observer's parameter in the Huggins detection function, that is,  $\approx 30$  or more observations where the observer had an opportunity to either detect or miss the burro group.

Given the current configuration for survey crews, sometimes the back-seat observer tasked with recording data will not have his or her full attention turned to the search for animals. There are in-flight practices that can mitigate the tendency for the data recorder to have lower average detection probabilities than he or she would, if he or she were not recording data. First, other crew members who observe burro groups should call out the observed covariates as quickly and clearly as possible so that the recorder can efficiently note the observations and return to searching for animals. Second, when passing or circling an observed group, data recorders should feel free to ask for more time to catch up on data recording before the pilot resumes forward progress along the transect. Even with these practices, the data recorder's attention can sometimes be focused on the data sheet when multiple groups are detected at the same time. It is important to have the same person work in this role as much as possible so that Huggins method models can accurately parameterize that person's acuity, given their additional data recording duties. In each of the data sets analyzed here, a single person recorded data.

### **Sample Size Considerations**

With any data set in which there is a small number of observations, the L-P 1 and L-P 2 techniques and the Huggins analytical approach can all lead to wide confidence intervals and, potentially, unreliable estimates. This was most notable at Warm Springs Canyon HMA, where a small number of observed burros led to substantially higher population estimates for the L-P 2 technique and Huggins analysis than the L-P 1 technique (table 7). That the L-P 2 estimate was high is not surprising, given that 7 of 8 observed burro groups happened to be seen on the left side of the flight line; technique L-P 2 assumes that an equal number of animals were on the right side of the flight line as the number that is estimated for the left side. That the Huggins estimate is high can be partially explained by the fact that data from the Warm Springs Canyon HMA and McGee Mountain HMA were analyzed together, and that in the overall data set there was estimated to be a decreasing probability of observing small groups compared to large groups (fig. 2). Correspondingly, the Huggins analysis estimated that the 8 observed small groups at Warm Springs Canyon HMA represented 11.7 missed groups, or 14.2 missed burros.

Having a large sample size of observations improves the precision of L-P and Huggins methods estimates. In prior analyses, the only source of uncertainty incorporated into the L-P 1 and L-P 2 estimators has been the binomial variance estimate for detection probability; this variance decreases as sample size increases. The Huggins analyses accounts for uncertainty in the estimated detection probability intercepts and in the estimated parameters representing each sighting covariate effect included in models. This uncertainty increases if there are few observations that can be used to parameterize each intercept or parameter. When sample size is small, the resulting, potentially high, uncertainty in the number of missed groups can translate to a high level of uncertainty in the number of missed animals for Huggins analyses. An alternate approach to estimating variance for Huggins models is to fit bootstrapped data to the candidate set of models (Wong, 1996; Lubow and Ransom, 2007; Griffin and others, 2013), but this approach also leads to high confidence intervals when sample sizes are low.

An advantage of Huggins analyses is that, as long as the survey protocols (including airspeed, height above ground, and season), some of the observers, and the type of helicopter used in a given HMA are consistent into the future, future Huggins analyses of abundance can use double-observer burro observation data from past surveys to increase the sample size of observations fit to models. Compared to the L-P techniques, it is a substantial benefit to the Huggins analyses that data sets from multiple survey years and survey sites can be combined to improve the precision of population abundance estimates. Whereas data sets can also be pooled for L-P analyses, those techniques do not allow for specific estimates of individual observer acuities. As a result, pooled L-P analyses would lead to detection probabilities averaged across multiple observers, presumably with confidence intervals for detection probability wide enough to encompass the process variance represented by variation in observer acuity. Insofar as Huggins type analyses can parameterize individual acuity, population size estimates that come from that technique may be expected to be more precise and accurate.

### **Effects of Transect Width**

Prior analytical methods in Arizona have used only those burros that were observed within 200 m of the transect line and have applied an area correction factor of 2, such that the number of burros estimated for that narrow transect is doubled to represent the entire population. This approach reflects an assumption that burros are uniformly distributed with respect to the helicopter flight path, and that they do not move either toward or away from the approaching helicopter. I question whether it is the best practice for analyses to ignore the observations made beyond the 200-m distance.

There were three data sets from Arizona for which I could compare L-P estimated population sizes based only on the burro groups within 200 m of the transect line to L-P estimated population sizes based on the entire set of burro observations (tables 4 and 7). Using only the smaller sample of observed burro groups close to the transect line necessarily limited the number of observations available to parameterize the L-P estimates of observer detection probability. At Lake Pleasant HMA, when only the observations within 200 m of the transect line were used, there was a sizable difference between the point estimates for burro abundance from L-P 1 ( $\hat{N}_{L-P 1} = 428$  burros) and L-P 2 ( $\hat{N}_{L-P 2} = 257$  burros). These differences may be indicative of how the L-P techniques can be sensitive to small sample sizes and variation in the side of the flight line where the burro groups were observed. With reference to all available burro observations at Lake Pleasant HMA, there was less of a difference between population size estimates for L-P 1 ( $\hat{N}_{L-P 1} = 350$  burros) and L-P 2 ( $\hat{N}_{L-P 2} = 306$  burros).

In many cases, there was a reasonable concordance between the population estimate from the Huggins analysis and the L-P 1 and L-P 2 analyses made with all available observations. The largest discrepancy in estimated population sizes was for the Black Mountain HMA fall data set in comparing the Huggins estimate (1,221 burros, 90 percent confidence interval (CI) = {1,170, 1,313}) to the L-P 1 and L-P 2 estimates that were based only on burro groups observed within 200 m of the transect (L-P 1 estimate = 2,145 burros, 90 percent CI = {1,842, 2,567}; L-P 2 estimate = 2,361 burros, 90 percent CI = {2,250, 2,484}). In that data set only, the frequency distribution of observed distances (fig. 1) was skewed to burro groups within 200 m of the transect line. Applying a correction factor of 2 to the L-P 1 and L-P 2 estimates made with reference to only those <200-m-distant burro groups led to burro population estimates that were far greater than those from the Huggins analysis or from the L-P 1 and L-P 2 analyses based on all available data. In some habitats, it is possible that detection distance in some data sets is so limited by distance that it would be most appropriate to only analyze observations made within a threshold distance, as has been done in Arizona in the past. Testing for the most accurate analysis method would best be done if there are pre-gather and post-gather surveys before and after a gather and removal of a substantial known number of burros or if there is a substantial number of radio-collared burros in a population. The 2014 surveys had neither removals nor radio-collared burros, so it is not possible, given these data, to know which of the population estimates for Black Mountain HMA is the most accurate.

The Huggins analyses for the effects of distance on detection probability (fig. 4) seems to suggest in all data sets analyzed here that observer detection does not fall off precipitously above a 200-m threshold. I interpret that result as an indication that any effects of distance can be modeled, and compensated for, by applying group-specific correction factors. One motivation for limiting observations to a narrow transect width is to avoid potentially strong effects of distance or vegetative cover on detection probability, but I suggest that a 200-m threshold may be too small, at least for the areas analyzed here. I did not find evidence of a dramatic decrease in detection probability beyond 200 m; also, limiting observations to those within 200 m of the transect line reduces the number of observations available for parameterizing detection probability functions. I acknowledge that the apparent lack of substantial decrease in detection probability as a function of distance in all data sets analyzed here could have been a result of some unmodeled heterogeneity bias manifested such that both front and back observers may have been likely to both see, or to both miss, burro groups at increasing distances. If this possibility were the case, though, I would also expect to see a sizable correlation between group size and distance because large far-away groups would be expected to be seen by both front and back observers more commonly than small far-away groups. I did not, however, detect a substantial correlation between group size and distance in any of the four data sets in which distance was recorded in 100-m increments. In the future, unmodeled heterogeneity bias could potentially be estimated if there are radio-marked animals in the surveyed population (Griffin and others, 2013) or if there is some other source of information about detection probabilities such as with the use of mark-recapture distance analyses (Burt and others, 2014).

I also note that the distribution of observed distances (fig. 1) is not entirely consistent with a no-movement assumption. In most cases the most commonly recorded distance was not the closest category; this may be indicative of burro movement away from the approaching helicopter or may simply be a result of an incomplete field of view within a short distance from the flight path, especially for the back-seat observers. What is clear is that all mark-recapture analyses can be very sensitive to small sample sizes. In the case of Lake Pleasant HMA, the

sample size of observations within 200 m of the flight path was quite limited ( $n=26$ ), such that the L-P 1 and L-P 2 analysis methods that applied the area correction factor of 2 yielded quite different results. I suggest that the use of the Huggins modeling approach, making use of all available data, is preferable to using only the limited data set recorded within 200 m of the flight path, at least when habitat conditions are similar to those that led to data sets analyzed in this report.

## Conclusions

It is unsatisfying to report estimated population sizes that differ so much depending on the analytical method chosen and, for Arizona, the choice of whether to include all observed burros or only those within 200 m of the transect line. I have presented some theoretical reasons to suggest that the Huggins analytical method may be preferable to the L-P 1 or L-P 2 techniques. I will leave it to others to compare the merits of the L-P 1 versus L-P 2 techniques, but I note again that neither of those methods makes use of observations that the pilot may contribute. Ultimately, though, it will not be possible to compare directly the accuracy of the analytical methods used in this report until there are areas where the true populations of burros in the surveyed areas are known, or until there are additional sources of data, such as a large-scale burro gather or radio-telemetry data, that can be used in analyses.

One method that has been used in the past to assess the accuracy of aerial surveys for wild horses and burros is to conduct a thorough pre-gather aerial survey, remove animals from the population with a gather, and to conduct a thorough post-gather aerial survey. This was the approach used in initial tests of the applicability of the double-observer method in Arizona for burro aerial surveys (BLM, written commun., 2000a, 2000b). When there are such opportunities in the future, I urge BLM to conduct pre-gather and post-gather aerial surveys, so that the accuracy of L-P 1, L-P 2, and Huggins methods can be compared.

On theoretical grounds, the Huggins estimator remains an attractive method for analyses of simultaneous double-observer data for wild burro surveys because of the ability to account for effects of sighting covariates such as group size, distance, and individual observer acuity, and because surveys conducted over multiple years can be used to overcome the analytical limitations that come from small sample sizes. Even after accounting for sighting covariates, it is likely that some fraction of the burro population remains unaccounted for, as a result of heterogeneity bias—this means that population estimates even from Huggins models should be viewed as likely to be underestimates of the true abundance of burros in the surveyed areas. Heterogeneity bias for HMAs of different vegetative and topographic types could be quantified if there were populations of burros with radio collars in those types or which were closely monitored by many teams of ground observers during aerial surveys (Griffin and others, 2013). Subsequent burro population estimates that use a heterogeneity bias parameter as part of Huggins analyses would then be expected to provide more accurate estimates of burro population size. I encourage the development of such estimates of heterogeneity bias, along with the testing of other potential techniques that may improve burro population estimation, such as the use of the latest high-definition infrared sensors.

## Acknowledgments

I gratefully acknowledge the contributions of the BLM Wild Horse and Burro specialists who organized and collected the data that were analyzed in this report: Chad Benson (Kingman Field Office, Arizona), Steve Bird (Hassayampa Field Office, Arizona), Dusty Carpenter (Vernal

Field Office, Utah), John Hall (Yuma Field Office, Arizona), Steve Surian (Sunrise Field Office, California), Garrett Swisher (Black Rock Field Office, Nevada), and Mike Tweddell (Price Field Office, Utah). I also thank the other staff that contributed to the success of the flights in their roles as State leads, observers, helicopter managers, dispatchers, and safety plan reviewers. I am very grateful to the pilots who capably flew the surveys. I am also grateful to Dave Conrad of Arizona Game and Fish Department and Dr. Bruce Lubow for use of early versions of spreadsheet templates that contributed to the analysis, to Gregor Auble for presubmission reviews, and to Dave Conrad, Larisa Harding, John Hervert, and Esther Rubin for peer reviews of the manuscript.

## References Cited

- Anderson, D.R., 2001, The need to get the basics right in wildlife field studies: *Wildlife Society Bulletin*, v. 29, p. 1294–1297.
- Barker, R., 2008, Theory and application of mark-recapture and related techniques to aerial surveys of wildlife: *Wildlife Research*, v. 35, p. 268–274.
- Beaver, J.T., Harper, C.A., Kissell, R.E., Muller, L.I., Basinger, P.S., Gode, M.J., Van Manen, F.T., Winton, W., and Kennedy, M.L., 2014, Aerial vertical-looking infrared imagery to evaluate bias of distance sampling techniques for white-tailed deer: *Wildlife Society Bulletin* v. 38, p. 419–427.
- Borchers, D.L., Buckland, S.T., Goedhart, P.W., Clarke, E.D., and Hedley, S.L., 1998, Horvitz-Thompson estimators for double-platform line transect surveys: *Biometrics*, v. 54, p. 1221–1237.
- Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., and Thomas, L., 2001, *Introduction to distance sampling*: Oxford, UK, Oxford University Press. 448 p.
- Bureau of Land Management (BLM), 1980, *Cibola-Trigo Herd Management Area plan—Final*: U.S. Department of Interior, Bureau of Land Management.
- Bureau of Land Management (BLM), 2010, *Wild horse and burro population inventory and estimation*: Washington, D.C., Bureau of Land Management Instructional Memorandum No. 2010-057, 4 p.
- Burnham, K.P., and Anderson, D.R., 2002, *Model selection and multimodel inference—A practical information-theoretic approach* (2d ed.): New York, Springer-Verlag, 353 p.
- Burt, M.L., Borchers, D.L., Jenkins, K.J., and Marques, T.A., 2014, Using mark-recapture distance sampling methods on line transect surveys: *Methods in Ecology and Evolution*, v.5, p. 1180–1191.
- Caughley, G., and Grice, D., 1982, A correction factor for counting emus from the air and its application to counts in Western Australia: *Australian Wildlife Research*, v. 9, p. 253–259.
- Graham, A., and Bell, R., 1989, Investigating observer bias in aerial survey by simultaneous double-counts: *Journal of Wildlife Management*, v. 53, p. 1009–1016.
- Griffin, P.C., Jenkins, K.J., Happe, P., Reid, M., Vales, D., Moeller, B.J., Reid, M., Happe, P.J., McCorquodale, S., Tirhi, M., Schaberl, J.P., and Beirne, K., 2013, A hybrid double-observer sightability model for aerial surveys: *Journal of Wildlife Management*, v. 77, p. 1532–1544.
- Griffin, P.C., Schoenecker, K.A., Gogan, P.J.P., and Lubow, B.C., 2009, *March 2009 ungulate survey on Santa Rosa Island, Channel Islands National Park*: U.S. Geological Survey Open-File Report 2009–1243, 23 p.
- Huggins, R.M., 1989, On the statistical analysis of capture experiments: *Biometrika*, v. 76, p. 133–140.

- Huggins, R.M., 1991, Some practical aspects of a conditional likelihood approach to capture experiments: *Biometrics*, v. 47, p. 725–732.
- Laake, J.L., Collier, B.A., Morrison, M.L., and Wilkins, R.N., 2011, Point-based mark-recapture distance sampling: *Journal of Agricultural, Biological, and Environmental Statistics*, v. 16, p. 389–408.
- Lubow, B.C., and Ransom, J.I., 2007, Aerial population estimates of wild horses (*Equus caballus*) in the Adobe Town and Salt Wells Creek Herd Management Areas using an integrated simultaneous double-count and sightability bias correction technique: U.S. Geological Survey Open-File Report 2007–127, 13 p.
- National Research Council, 2013, Using science to improve the BLM wild horse and burro program—A way forward: Washington, D.C., The National Academies Press, 383 p.
- Pollock, K.H., and Kendall, W.L., 1987, Visibility bias in aerial surveys—A review of estimation procedures: *Journal of Wildlife Management*, v. 51, p. 502–510.
- Ransom, J., 2012, Detection probability in aerial surveys of feral horses: *Journal of Wildlife Management*, v. 76, p. 299–307.
- Seber, G.A.F., 1982, The estimation of animal abundance and related parameters (2d ed.): New York, Macmillan, 654 p.
- Southwell, C., Borchers, D., Paxton, C.G.M., Burt, L., and de la Mare, W., 2007, Estimation of detection probability in aerial surveys of Antarctic pack-ice seals: *Journal of Agricultural, Biological, and Environmental Statistics*, v. 12, p. 41–54.
- Steinhorst, R.K., and Samuel, M.D., 1989, Sightability adjustment methods for aerial surveys of wildlife populations: *Biometrics*, v. 45, p. 415–425.
- Walter, M.J., and Hone, J., 2003, A comparison of 3 aerial survey techniques to estimate wild horse abundance in the Australian Alps: *Wildlife Society Bulletin*, v. 31, p. 1138–1149.
- White, G.C., and Burnham, K.P., 1999, Program MARK—Survival estimations from populations of marked animals: *Bird Study*, v. 46 (Supplement), p. 120–138.
- Wong, C., 1996, Population size estimation using the modified Horvitz-Thompson estimator with estimated sighting probabilities: Fort Collins, Colorado State University, Ph.D. dissertation, 202 p.

**Appendix.** Sighting covariates for each observed burro group. Text in bold indicates choices that could have been recorded. For any given sighting covariate, only one choice was recorded.

[m, meter]

Sighting covariate	Description
Total count	Number of %burros in the detected group. This is recorded as adults and foals.
Side of ship	Location of group when first detected, relative to the flight line. <b>L</b> = left of transect; <b>R</b> = right of transect; <b>B</b> = both (that is, for large groups that span the transect and <i>could</i> have been seen from both rear seats, or a group that runs from one side to the other under aircraft and was available to be detected by both sides of the aircraft at some point); and <b>Center only</b> = the group was never visible to the rear seats because it was always right under the helicopter's flight line.
Activity	Activity level of the burro group when it was first detected. <b>Move</b> = moving; <b>Still</b> = still when you first saw the horses or burros.
Visual field	The one <i>dominant</i> type of background in the visual field near the burro group, including an area up to 10 m from the group when first detected. <b>Open</b> = 75 percent of the ground is mainly grass, bare ground, vegetation below the height of burros' hocks; <b>Broken</b> = 25 percent or more of the ground is covered by vegetation from hocks to burro-head high; <b>Tree</b> = 25 percent or more of the ground-covering vegetation is taller than burro head. If tall shrubs and trees both cover, 25 percent or more of the ground, choose <b>Tree</b> .
Percent concealing vegetation	The percentage of the burro-height space obscured from the observers' view by vegetation cover, in the entire area of the group when first detected, including up to 10 m away from the entire group (all of the included individuals). "Concealing Vegetation" means plant material (stem, branch, foliage) of any combination of plant species that obscure the burro-height space. This is approximated and recorded in increments to the nearest 10 percent.
Distance	The approximate horizontal distance from the burro group to the nearest transect. Recorded in most surveys as < <b>100 m</b> , <b>100–200 m</b> , <b>200–300 m</b> , <b>300–400 m</b> , or <b>&gt;400 m</b> . In some surveys, this was recorded as <¼ mile, ¼ to ½ mile, or ½ to 1 mile.
Light level	The light-level condition in the immediate vicinity of the burro group. <b>Hi</b> = high contrast; there was bright light and visible shadows at the burro group. Burros in tree shade are recorded as "Hi." <b>Flat</b> = flat light; the entire area between aircraft and burro group was in shadow of overhead clouds, with no distinct shadows; <b>Shade</b> = the burro group was in the shade from a particular cloud, cliff, or rock, but there were sunny places between the aircraft and the burros (so the burros were in the dark, but there was light ground in the intervening space).
Topography class	<b>Smooth</b> = open terrain, little to no topography or rocks that could hide a burro. <b>Rugged</b> = rugged terrain with rocks or topography tall enough to hide a horse or burro. This is assessed at the scale of an oval up to 10 m (10 yards) beyond the edge of the whole group.



