To: Alan Shepherd (BLM), Amy Dumas (BLM), Robert Hopper (BLM)
From: Bruce Lubow (CSU), Jason Ransom (USGS), Gail Collins (USFWS)
Date: 18 August 2010
RE: Memo on Tri-State Horse Survey Results

*Note: *This memo is intended for internal use by the addressee(s) and to provide information specific to management needs. It has not been peer-reviewed and is not considered published literature by the USGS.*

Summary

An extensive aerial survey of wild horse populations in 20 Federal management units in the tri-state areas of California, Oregon, and Nevada was conducted in June 2010 by USFWS and BLM personnel. Data were analyzed by USGS to provide statistical estimates of sighting probabilities as a function of individual observers, horse group size, horse activity, distance of horses from the observers, vegetation characteristics, and terrain type. These estimates were then used to correct the raw counts for systematic biases (undercounts) that inevitably occur on wildlife surveys. The statistical methods also provided estimates of the likely error (uncertainty) in measuring the size of this population. Sighting probabilities were generally very high; overall, 97.7% of the estimated population was seen by at least one of the 4 aerial observers. Estimates of the number of horses present in the various management units at the time of the survey ranged from as few as 3 to as many as 1,236. The combined estimate for the entire area surveyed was 4,315 (95% CI = 4,234-4,424) horses. Although some unquantifiable errors could remain, these are most likely to have resulted in our estimate being too low rather than too high. The excellent sighting conditions and resulting high sighting probabilities strongly support the use of these estimates as a basis for management decisions.

Field methods

A series of flights was conducted on 9 days over a 10 day period from June 20 to June 29, 2010 to survey mostly contiguous Federal Lands (USFWS and BLM) in the tri-State area surrounding the common borders of California, Oregon, and Nevada (no flights were conducted on June 25). The survey was flown in a fixed-wing airplane (Cessna 210 Turbo) along parallel transects running north and south, spaced 1.5 miles apart. To attempt to minimize the potential for horse movement between survey days, where possible, barriers to movement (fences and terrain) were used as break-off points to divide the larger survey area. A pilot and 3 additional

observers from USFWS and BLM independently recorded groups of horses when sighted, along with the number of horses in each group and several additional variables that were expected to affect the probability of sighting a group. These variables included the terrain type (rugged or not), distance from the aircraft (0-0.25 mile, 0.25-0.5 mile, 0.5-1.0 mile, or 1.0-1.5 miles), cover type (open grass or other unobstructed versus trees or shrubs), percentage of the area where horses were located that was covered by potentially obstructing vegetation (shrubs or trees), and horse movement (moving or still). Groups were recorded by position relative to the flight line, either left, right, center, or both (spread across the flight line and visible to observers on both sides).

Each of the 4 observers independently searched for horses. When necessary, observers could announce the sighting of a group to all members of the team so that the plane could circle back to allow for a more careful count of the group size and recording of covariates. Presence of a group was announced only after the plane had fully passed a beam of the group giving all observers a chance to see it independently. Observers who had not seen a group prior to an announcement recorded the group as missed (not seen). In all cases, a single set of covariates was determined for each group by consensus of the observers—only the initial success or failure in sighting a group was determined independently for each observer. The same pilot (left side) and two observers on the right side (front and back) remained in the same positions on all flights. Three additional observers took part in portions of the survey, taking turns over the several days in the back-left seat.

Analysis Methods

We used the Huggins closed capture model structure in Program MARK to fit multiple plausible models of sighting probability to the data. This model enables inclusion of covariates that affect sighting probability of individual groups as well as accounting for differences among the observers of the survey. The model of sighting probability for group *i* is a logistic function of the following form:

$$p_i = \frac{1}{1 + e^{-(\hat{\beta}_{obs} + \hat{\beta}_N N + \hat{\beta}_T T_i + \hat{\beta}_A A + \hat{\beta}_B B + \hat{\beta}_C C_i + \hat{\beta}_D D_i + \hat{\beta}_M M_i + \hat{\beta}_V V_i)}}$$

Where:

$\hat{\beta}_{obs}$=Estimated intercept for a specific observer and group position combination (see table 1)

$\hat{\beta}_N$ = estimated slope for group size effect

$\hat{\beta}_T$ = estimated slope for terrain effect

$\hat{\beta}_A$ = estimated slope for first back-left observer

$\hat{\beta}_B$ = estimated slope for second back-left observer

$\hat{\beta}_C$ = estimated slope for cover type effect
$\hat{\beta}_D$ = estimated slope for distance effect
$\hat{\beta}_M$ = estimated slope for movement effect
$\hat{\beta}_V$ = estimated slope for vegetation effect

N = group size (no. horses)
T = terrain type (rugged=1, other=0)
A = first back-left observer (1 if observer A present, 0 otherwise)
B = second back-left observer (1 if observer B present, 0 otherwise)
C = cover type (open=1, other=0)
D = distance to mid-point of distance category: 0.125, 0.375, 0.75, or 1.25 miles.
M = movement (moving=1, still=0)
V = proportion of ground covered in obscuring vegetation (shrubs or trees)

(Some additional model structures were considered, but received minimal support. These are detailed in Table 2).

The rear observers had no ability to see horses along the center of the transect line (directly in front of the plane). Similarly, the front-right observer was too short to easily see over the console and consequently did not attempt to survey the centerline. The pilot scanned both sides of the flight line for other aircraft as well as horses, whereas the other observers focused strictly on surveying their own side. (One instance of a rear observer sighting a group on the opposite side of the plane was an anomaly and was excluded from the analysis.) Sighting probabilities were assumed to be the same for groups recorded as being on the same side as an observer or on both sides of the flight line. We assumed that the pilot had equal sighting probability for groups on the same side and on the center line because there were too few observations on the center line (2) to estimate a separate probability and because no other observers could see the centerline and we thus had no way to estimate the pilot's ability there. Given these considerations, the parameters we included for $\hat{\beta}_{obs}$ were as follows:

Table 1. Intercept parameters of sightability model for various observer, horse position combinations.

	Horse Group Position			
Observer Position	Left	Right	Both	Center
Front-left (pilot)	β_1	β_2	β_1	β_1
Front-right	0	β_3	β_3	0
Back-left	β_4	0	β_4	0
Back-right	0	β_5	β_5	0

Thus, $\hat{\beta}_{obs}$ represents these estimated 5 parameters. We also considered one simplified model with rear observers having equal probability ($\beta_4 = \beta_5$), thereby reducing the observer parameters to 4. We designate these alternatives as $\hat{\beta}_{obs,5}$ and $\hat{\beta}_{obs,4}$. We did not consider models with either pilot or front-right observer probabilities set equal to any other values, because preliminary analyses indicated that such models received essentially no support from the data. We considered 35 models with various combinations of the observer and covariate parameters (Table 2).

To fully account for all of the measurable sources of uncertainty, we applied a bootstrap procedure. We constructed an empirical population containing one group with identical observation values to the actual groups seen for each whole group estimated from that observation, plus an additional group with a probability equal to the fractional portion of the estimated number of groups. In other words, for a group with a sighting probability of 0.4, we would estimate 1/0.4 = 2.5 groups. The empirical population would contain 2 groups with values identical to the one actually observed, plus a third identical group with 50% probability in each bootstrap simulation.

We then simulated the survey of each empirical population by randomly determining which observers saw each group, based on the original estimates of their group-specific probability of sighting that group. Next, we refit all of the mark-resighting models in the inference set (those with QAICc < 4.0 in the original analysis) to the simulated survey data and computed population estimates from the simulated data with the sighting models refit to these simulated data and using the model weights calculated from this refitting to the simulated data. We repeated this process (simulate an empirical population, simulate a survey of this population, fit models to the simulated data, and compute weighted estimates) for 99 simulation iterations and recorded the population estimates from each. We calculated standard errors for the original population estimates as the standard deviation of the 100 estimates (99 simulated plus 1 original) and the 95% confidence intervals as the 2.5-percentile and 97.5-percentile of the simulated values. These error measures fully account for binomial sampling error, sighting model estimation error, covariances in sighting probabilities, model selection error, and extra-binomial variation (lack of model fit).

Results

The variance inflation factor for the most general model was 1.87, so we made inferences using quasi-likelihood methods. Of the 35 models considered, 17 had values with ΔQAICc <4.0. The effect of group size was overwhelmingly supported (QAICc evidence ratio 388:1) and, as expected, sighting probability increased rapidly with group size. There was stronger support (QAICc evidence ratio 2.73:1) for the $\hat{\beta}_{obs,4}$ intercepts (left and right rear observers had equal sighting probability) than for $\hat{\beta}_{obs,5}$ (unique sighting probabilities for back left and right observers). Terrain type was also well supported (QAICc evidence ratio 1.92). We included these 17 in the inference set for further analysis. Other effects were weakly supported (QAICc

evidence ratio < 0.5:1). The number of horses for each management unit was estimated using the QAICc weighted estimates from all 17 models in this set, which takes into account all of the various effects of covariates to the extent that the weight of evidence supports each effect.

Based on the most strongly supported model, sighting probabilities were high for all primary observers and adequate for the pilot. For a horse on the same side as the observer under worst-case sighting conditions (a single horse, group size = 1, in rugged terrain), the front right observer had an estimated sighting probability of 64.9%, the rear observers had 82.5%, and the pilot had 26.5%. These values rose to 77.5%, 89.7% and 40.2%, respectively, for a single horse in even terrain. For a typical group (~10 horses) these sighting probabilities (in even terrain) rose to 92.3%, 82.5% and 47.8%, respectively. Sighting probabilities rose rapidly with group size and were essentially 100% for all observers for the largest group (188 horses). Bear in mind that the probability of missing a group is the combined probability of both front and rear observers missing it. So, in the case of a group of 10 in even terrain, the sighting probability is 1 – (1-92.3%)·(1- 82.5%) = 98.7% on the right side and 1 – (1-92.3%)·(1- 47.8%) = 96.0% on the left side.

The estimated population for all 20 management units combined was 4,315 (95% CI = 4,234-4,424) horses and ranged from as few a 3 to as many as 1,225 horses in individual management units (Table 3). The percentage of horses seen by observers was 97.7% overall and ranged from 94.0% to 99.2% across the 20 management units. Coefficient of variation (SE/Estimate) was only 1.2% overall and ranged from 1.0% to 38.8% across management units, with the largest error estimates occurring in units with very few observed horses.

Many of the lower confidence interval limits are actually less than the number of horses sighted during the survey. For example, the lower 95% confidence limit for the Sheldon NWR is 1,215 even though 1,225 horses were actually seen. This is a normal statistical result and reflects the fact that a confidence interval expresses what would likely happen if the survey were repeated. In 95% of repeated surveys, the estimate would be >1,215. Some surveys would miss more horses and produce lower estimates, even after corrections, than this survey actually did. For management purposes, it is reasonable to assume that there are at least as many horses as were observed during this survey, rather than using the lower confidence limit as a minimum number.

Discussion

The reliability of these results rests on several important assumptions. It is assumed that all horse groups are flown over one time during the survey period and thus have exactly one chance to be counted. Although attempts were made to minimize the potential for horse movement between survey days, if horses were to move between days of the survey so as to be present and available to observers more than once or not at all, this could potentially bias the results. However, if such movements were random, so that it was just as likely that a horse group was available for observation twice as unavailable at any time, the results would not be biased. Only systematic movements would pose a potentially uncorrectable error. It is also assumed that all horse groups with identical sighting covariate values have equal sighting probability. If there is additional variability in sighting probability not accounted for in the sighting models, such heterogeneity would lead to a negative bias (underestimate) of the population. With sighting conditions as good as they were in this survey, this problem is less likely. The variance inflation factor indicates some minor lack of fit in the model, possibly due to heterogeneity of sighting probabilities, but a value of 1.87 is only slightly elevated (1.0 is ideal and values >3 are considered serious). A third potential bias is created if the number of horses in each group is not counted accurately. This is particularly likely to occur in very large groups where it is common to miss a few horses unless photographs or video is used and scrutinized following the flight. Given the various potential sources of bias, it is more likely that our estimates are somewhat lower, rather than higher, than the true population. Given the very high sighting probabilities estimated for this survey, we are confident that our estimates provide a sound and reliable basis for management decisions.

Table 2. Results of sighting probability model fitting for 9-day survey of 20 wild horse management units in California, Oregon, and Nevada in June 2010. Models in bold have $\Delta QAIC_c < 4.0$ and were included in the inference set.

	Model	$QAIC_c$	Delta $QAIC_c$	$QAIC_c$ Weights	Model Quasi-Likelihood	Num. Par	Q-Deviance
1	$\boldsymbol{\hat{\beta}_{obs,4}+\hat{\beta}_N}$**N+**$\boldsymbol{\hat{\beta}_T}$**T**	**369.5510**	**0**	**0.16477**	**1**	**6**	**357.4940**
2	$\boldsymbol{\hat{\beta}_{obs,4}+\hat{\beta}_N}$**N**	**370.8652**	**1.3142**	**0.08541**	**0.5184**	**5**	**360.8245**
3	$\boldsymbol{\hat{\beta}_{obs,4}+\hat{\beta}_N}$**N+**$\boldsymbol{\hat{\beta}_T}$**T+**$\boldsymbol{\hat{\beta}_D}$**D**	**371.0503**	**1.4993**	**0.07786**	**0.4725**	**7**	**356.9742**
4	$\boldsymbol{\hat{\beta}_{obs,4}+\hat{\beta}_N}$**N+**$\boldsymbol{\hat{\beta}_T}$**T+**$\boldsymbol{\hat{\beta}_C}$**C**	**371.2149**	**1.6639**	**0.07171**	**0.4352**	**7**	**357.1388**
5	$\boldsymbol{\hat{\beta}_{obs,4}+\hat{\beta}_N}$**N+**$\boldsymbol{\hat{\beta}_T}$**T+**$\boldsymbol{\hat{\beta}_M}$**M**	**371.4300**	**1.8790**	**0.06439**	**0.3908**	**7**	**357.3539**
6	$\boldsymbol{\hat{\beta}_{obs,4}+\hat{\beta}_N}$**N+**$\boldsymbol{\hat{\beta}_T}$**T+V**	**371.4942**	**1.9432**	**0.06236**	**0.3785**	**7**	**357.4181**
7	$\boldsymbol{\hat{\beta}_{obs,5}+\hat{\beta}_N}$**N+**$\boldsymbol{\hat{\beta}_T}$**T**	**371.5566**	**2.0056**	**0.06044**	**0.3668**	**7**	**357.4805**
8	$\boldsymbol{\hat{\beta}_{obs,4}+\hat{\beta}_N}$**N+**$\boldsymbol{\hat{\beta}_D}$**D**	**372.1293**	**2.5783**	**0.04539**	**0.2755**	**6**	**360.0722**
9	$\boldsymbol{\hat{\beta}_{obs,4}+\hat{\beta}_N}$**N+**$\boldsymbol{\hat{\beta}_C}$**C**	**372.7431**	**3.1921**	**0.0334**	**0.2027**	**6**	**360.6861**
10	$\boldsymbol{\hat{\beta}_{obs,4}+\hat{\beta}_N}$**N+**$\boldsymbol{\hat{\beta}_M}$**M**	**372.7949**	**3.2439**	**0.03254**	**0.1975**	**6**	**360.7378**
11	$\boldsymbol{\hat{\beta}_{obs,5}+\hat{\beta}_N}$**N**	**372.8441**	**3.2931**	**0.03175**	**0.1927**	**6**	**360.7871**
12	$\boldsymbol{\hat{\beta}_{obs,4}+\hat{\beta}_N}$**N+V**	**372.8699**	**3.3189**	**0.03135**	**0.1903**	**6**	**360.8129**
13	$\boldsymbol{\hat{\beta}_{obs,5}+\hat{\beta}_N}$**N+**$\boldsymbol{\hat{\beta}_T}$**T+**$\boldsymbol{\hat{\beta}_D}$**D**	**373.0647**	**3.5137**	**0.02844**	**0.1726**	**8**	**356.9668**
14	$\boldsymbol{\hat{\beta}_{obs,4}+\hat{\beta}_N}$**N+**$\boldsymbol{\hat{\beta}_T}$**T+**$\boldsymbol{\hat{\beta}_M A+\hat{\beta}_M B}$	**373.1415**	**3.5905**	**0.02737**	**0.1661**	**8**	**357.0436**
15	$\boldsymbol{\hat{\beta}_{obs,5}+\hat{\beta}_N}$**N+**$\boldsymbol{\hat{\beta}_T}$**T+**$\boldsymbol{\hat{\beta}_C}$**C**	**373.2287**	**3.6777**	**0.02620**	**0.1590**	**8**	**357.1308**
16	$\boldsymbol{\hat{\beta}_{obs,5}+\hat{\beta}_N}$**N+**$\boldsymbol{\hat{\beta}_T}$**T+**$\boldsymbol{\hat{\beta}_M}$**M**	**373.4373**	**3.8863**	**0.02360**	**0.1432**	**8**	**357.3394**
17	$\boldsymbol{\hat{\beta}_{obs,5}+\hat{\beta}_N}$**N+**$\boldsymbol{\hat{\beta}_T}$**T+V**	**373.5034**	**3.9524**	**0.02284**	**0.1386**	**8**	**357.4055**
18	$\hat{\beta}_{obs,5}+\hat{\beta}_N$N+$\hat{\beta}_D$D	374.1240	4.5730	0.01674	0.1016	7	360.0480
19	$\hat{\beta}_{obs,4}+\hat{\beta}_N$N+$\hat{\beta}_M A+\hat{\beta}_M B$	374.3498	4.7988	0.01496	0.0908	7	360.2737
20	$\hat{\beta}_{obs,5}+\hat{\beta}_N$N+$\hat{\beta}_T$T+$\hat{\beta}_M A+\hat{\beta}_M B$	374.4602	4.9092	0.01415	0.0859	9	356.3378
21	$\hat{\beta}_{obs,5}+\hat{\beta}_N$N+$\hat{\beta}_C$C	374.7303	5.1793	0.01237	0.0751	7	360.6542
22	$\hat{\beta}_{obs,5}+\hat{\beta}_N$N+$\hat{\beta}_M$M	374.7749	5.2239	0.01209	0.0734	7	360.6988
23	$\hat{\beta}_{obs,5}+\hat{\beta}_N$N+V	374.8521	5.3011	0.01163	0.0706	7	360.7760
24	$\hat{\beta}_{obs,5}+\hat{\beta}_N$N+$\hat{\beta}_M A+\hat{\beta}_M B$	375.7978	6.2468	0.00725	0.0440	8	359.6999

25	$\hat{\beta}_{obs,4}$+$\hat{\beta}_T$T+N*Obs[1]	375.9190	6.3680	0.00682	0.0414	10	355.7692
26	$\hat{\beta}_{obs,4}$+N*Obs[1]	377.2882	7.7372	0.00344	0.0209	9	359.1658
27	$\hat{\beta}_{obs,4}$+$\hat{\beta}_N$N+Day(7)[2]	377.9243	8.3733	0.00250	0.0152	12	353.7116
28	$\hat{\beta}_{obs,5}$+$\hat{\beta}_T$T+N*Obs[1]	377.9416	8.3906	0.00248	0.0151	11	355.7618
29	$\hat{\beta}_{obs,4}$+$\hat{\beta}_N$N+$\hat{\beta}_T$T+Day(7)[2]	378.0333	8.4823	0.00237	0.0144	13	351.7850
30	$\hat{\beta}_{obs,5}$+$\hat{\beta}_N$N*Obs[1]	379.2938	9.7428	0.00126	0.0076	10	359.1441
31	$\hat{\beta}_{obs,5}$+$\hat{\beta}_N$N+ Day(7)[2]	379.9231	10.372	0.00092	0.0056	13	353.6748
32	$\hat{\beta}_{obs,5}$+$\hat{\beta}_N$N+$\hat{\beta}_T$T+Day(7)[2]	380.0548	10.504	0.00086	0.0052	14	351.7682
33	$\hat{\beta}_{obs,4}$	382.7523	13.201	0.00022	0.0013	4	374.7252
34	$\hat{\beta}_{obs,5}$	384.6818	15.131	0.00009	0.0005	5	374.6411
35	$\hat{\beta}_{obs,5}$+$\hat{\beta}_N$N+$\hat{\beta}_T$T+$\hat{\beta}_C$C+$\hat{\beta}_V$V+$\hat{\beta}_D$D+$\hat{\beta}_M$M+$\hat{\beta}_M A$+$\hat{\beta}_M B$+Day(7)[2]	387.9653	18.414	0.00002	0.0001	19	349.4448

[1] Abbreviated notation indicating separate group size slopes corresponding to each observer intercept.

[2] Abbreviated notation indicating slopes for each survey day except the first day (7 parameters).

Table 3. Population results of 9-day survey of 20 wild horse management units in California, Oregon, and Nevada in June 2010. Results are based on average across 17 models in inference set.

	Area	Seen	Estimate	SE[1]	CV[2]	LCL[3]	UCL[3]	%Seen
1	Burns -- Pueblo-Lone Mtn. HA	19	19	1.91	10.1%	19	19	99.2%
2	Lakeview -- Beaty's Butte HMA	377	386	14.52	3.8%	350	410	97.8%
3	Lakeview -- Outside HMA/HA	4	4	0.57	14.3%	4	4	98.0%
4	Surprise Valley -- Bitner HMA	47	48	3.42	7.1%	41	55	98.0%
5	Surprise Valley -- Carter Reservior HMA	6	6	1.54	25.2%	0	7	94.8%
6	Surprise Valley -- Fox Hog HMA	300	317	15.00	4.7%	294	358	94.5%
7	Surprise Valley -- High Rock HMA	300	303	11.48	3.8%	290	329	98.9%
8	Surprise Valley -- Massacre Lakes HMA	148	149	4.74	3.2%	139	158	99.6%
9	Surprise Valley -- New Year's Lake HA	100	102	4.67	4.6%	88	109	98.3%
10	Surprise Valley -- Nut Mountain HMA	3	3	1.22	38.8%	0	7	94.0%
11	Surprise Valley -- Outside HMA/HA	302	319	18.86	5.9%	288	362	94.7%
12	Surprise Valley -- Wall Canyon HMA	87	88	3.00	3.4%	81	94	99.3%
13	Winnemucca -- Black Rock East HMA[4]	69	71	5.64	10.7%	41	63	97.1%
14	Winnemucca -- Black Rock West HMA	140	143	11.38	7.9%	121	168	97.7%
15	Winnemucca -- Calico Mtns. HMA	436	450	15.33	3.4%	421	479	96.9%
16	Winnemucca -- Granite Range HMA	154	160	13.09	8.2%	133	185	96.4%
17	Winnemucca -- Outside HMA/HA	136	138	5.47	4.0%	126	149	98.5%
18	Winnemucca -- Warm Springs HMA	342	350	16.64	4.7%	312	387	97.9%
19	Hart Mtn. NAR -- Sheldon-Hart Mtn. NWRC	22	22	2.94	13.1%	14	31	98.4%
20	Sheldon NWR -- Sheldon-Hart Mtn. NWRC	1225	1236	12.85	1.0%	1,215	1,263	99.1%
	All Areas Combined[4]	**4,217**	**4,315[5]**	**50.26**	**1.2%**	**4,234**	**4,424**	97.7%

[1] Standard Error

[2] Coefficient of variation

[3] Upper and lower 95% confidence interval

[4] On group of 19 horses seen by right rear observer on left side of aircraft were included in seen and estimated values, but were not used in statistical analysis of sighting probability and no corrections were made for this group.

[5] Does not match sum of column due to rounding.