2014

Field-readable alphanumeric flags are valuable markers for shorebirds: use of double-marking to identify cases of misidentification

Erin A. Roche
USGS Northern Prairie Wildlife Research Center, eroche@usgs.gov

Colin M. Dovichin
USGS Northern Prairie Wildlife Research Center

Todd W. Arnold
University of Minnesota, Saint Paul, arnol065@umn.edu

J. Field Ornithol. 85(3):329–338, 2014 DOI: 10.1111/jofo.12072

Field-readable alphanumeric flags are valuable markers for shorebirds: use of double-marking to identify cases of misidentification

Erin A. Roche,[1,3] Colin M. Dovichin,[1] and Todd W. Arnold[2]

[1] *U.S. Geological Survey, Northern Prairie Wildlife Research Center, 8711 37th Street SE, Jamestown, North Dakota 58401, USA*

[2] *Department of Fisheries, Wildlife, and Conservation Biology, University of Minnesota, 135 Skok Hall, Saint Paul, Minnesota 55108, USA*

Received 11 February 2014; accepted 13 May 2014

ABSTRACT. Implicit assumptions for most mark-recapture studies are that individuals do not lose their markers and all observed markers are correctly recorded. If these assumptions are violated, e.g., due to loss or extreme wear of markers, estimates of population size and vital rates will be biased. Double-marking experiments have been widely used to estimate rates of marker loss and adjust for associated bias, and we extended this approach to estimate rates of recording errors. We double-marked 309 Piping Plovers (*Charadrius melodus*) with unique combinations of color bands and alphanumeric flags and used multi-state mark recapture models to estimate the frequency with which plovers were misidentified. Observers were twice as likely to read and report an invalid color-band combination (2.4% of the time) as an invalid alphanumeric code (1.0%). Observers failed to read matching band combinations or alphanumeric flag codes 4.5% of the time. Unlike previous band resighting studies, use of two resightable markers allowed us to identify when resighting errors resulted in reports of combinations or codes that were valid, but still incorrect; our results suggest this may be a largely unappreciated problem in mark-resight studies. Field-readable alphanumeric flags offer a promising auxiliary marker for identifying and potentially adjusting for false-positive resighting errors that may otherwise bias demographic estimates.

RESUMEN. **Banderas alfanuméricas legibles a campo son valiosas marcas para aves costeras: el uso de marcado doble para identificar casos de mala identificación**
Los supuestos implícitos en los estudios de recaptura son que los individuos no pierden las marcas y que todas las marcas observadas han sido registradas correctamente. Si estos supuestos son violados, e. g., debido a la pérdida o al desgaste extremo de los marcas, los estimadores de tamaño de población y de tasas de vida serán sesgados. Experimentos con marcado doble han sido ampliamente utilizados para estimar la tasa de pérdida de marcas y ajustar por el sesgo asociado, y nosotros hemos extendido este método a estimar la tasa de errores de registro de las mismas. Marcamos de manera doble a 309 Playeros Melódicos (*Charadrius melodus*) con una única combinación de bandas de color y banderas alfa numéricas y usamos modelos de recaptura multi estado para estimar la frecuencia en que los playeros fueron mal identificados. Los observadores tienen el doble de probabilidad de leer y reportar una combinación de bandas de color invalida (2.4% del tiempo) que un código alfanumérico invalido (1.0%). Los observadores fallaron en leer combinaciones de bandas de color o de banderas de códigos alfanuméricos 4.5% del tiempo. A diferencia de estudios previos de recaptura por observación de bandas, el uso de dos marcas nos permitió identificar cuando los errores de observación resultaban en reportes de combinaciones o códigos que eran válidos, pero incorrectos; nuestros resultados sugieren que este puede ser un problema subestimado en los estudios de recaptura visual de marcas. Las banderas alfanuméricas legibles a campo ofrecen un marcador auxiliar promisorio para identificar y potencialmente ajustar por falso positivo los errores de recaptura visual que de otra forma sesgarían los estimadores demográficos.

Key words: Charadrius melodus, color bands, double-marking, false positives, mark-recapture, misidentification, Piping Plover

Mark-recapture, where animals are uniquely marked and their fates followed, is a technique that has been used for more than a century to understand the demographics of populations (Petersen 1896, Lincoln 1930, Lebreton et al. 1992). Although theoretically a straightforward concept, important assumptions must be met for estimates to be reliable, and one such assumption is that the unique marks applied to the sample of individuals remain equally identifiable over time (Anderson et al. 1985, Pollock et al. 1990). At one time this simply meant that the markers remained intact and readable. If this assumption was violated, misidentification and ultimately

[3] Corresponding author. Email: eroche@usgs.gov

biased and imprecise demographic estimates resulted (Arnason and Mills 1981, Mills et al. 2000, Creel et al. 2003, Yoshizaki et al. 2009, Morrison et al. 2011).

Misidentification simultaneously leads to two kinds of error; a false-positive error occurs when the wrong animal is assumed to have been encountered because a marker was incorrectly observed, and a false-negative error occurs when the animal that was actually observed is presumed to have gone undetected. For example, in photo-identification studies, misidentification that does not match a previously identified individual would be incorrectly assigned as a first-time capture. In scenarios where an observed marker does not exactly match one previously applied to an animal, investigators likely would discard that observation, but they might be tempted to use logical criteria to try to ascertain which individual was observed, thus increasing the probability of making a false-positive error (e.g., "blue-orange-blue was never deployed, so I must have seen blue-red-blue"). However, in studies with large numbers of previously marked animals, the chances of matching a previously marked, but incorrect, individual will escalate. False positives seem especially problematic because they can suggest capture histories that never occurred (e.g., an individual not seen for several years and then returning from the dead, or a long-distance migration). False negatives seem more benign because they would be interpreted as detection failure, but in cases where an individual is consistently missed due to false-negative errors (e.g., due to badly faded color bands), such errors could lead to underestimation of survival. When animals are captured, there is only a small chance that markers will be misread or misreported. However, use of evolving natural markings in photo-identification studies (Stevick et al. 2001, Marshall and Pierce 2012), genetic mark-recapture (Lukacs and Burnham 2005a), and automated identification processes (Hastings et al. 2008) has led to an increased appreciation that identification error falls along a continuum.

Recent studies have reported methods for statistically accounting for misidentification errors caused by allelic drop-out in genetic mark-recapture studies (false-positive error, Wright et al. 2009), genetic "shadow effects" (false-positive and false-negative errors, Mills et al. 2000), evolving natural marks (false-negative error, Link et al. 2010), incorrectly detecting a species in a point count when it is not present (false-positive error, Royle and Link 2006), and band loss (false-negative error, Conn et al. 2004, Juillet et al. 2010). These models have applications beyond the specific instances for which they were developed both for estimating the degree of misidentification error and for adjusting the demographic estimates of interest. However, these models largely require the application of two marking methods, with at least one method considered either permanent or definitive, e.g., multiple natural markers (Marshall and Pierce 2012), genotyping and radio-telemetry (Lukacs and Burnham 2005b), tattoos (Diefenbach and Alt 1998), and observer experience (Miller et al. 2011).

The practice of double-marking has long been used in bird studies and is recommended for long-term demographic studies (Kendall et al. 2009); the most prominent example being the use of metal leg bands along with some other form of less-permanent auxiliary marker such as a plastic neck band, nasal marker, or colored leg bands (Nichols and Hines 1993, Zimmerman et al. 2009). For example, the use of unique combinations of plastic color bands has become increasingly popular for identifying individual birds with easily visible legs in the field without having to recapture individuals. However, despite its popularity, band loss (Ottaway et al. 1984, Spendelow et al. 1994), color fading (Ottaway et al. 1984), and resighting errors (Milligan et al. 2003) can make this marking technique problematic.

Historically, information gleaned from double-marking has been used to estimate retention rates of auxiliary markers and adjust demographic estimates of interest for the bias that results (Weiss et al. 1991, Kendall et al. 2009). Resighting errors have been studied less extensively and could similarly be addressed with the use of a double-marking strategy. Typically, investigators choose from a known pool of possible marker combinations, are aware of all combinations that have ever been used in their mark-recapture study, and can simply disregard reports of combinations that do not exist. However, as the number of used (and thus possible) band combinations increases, so too does the probability that what would once have been an easily discarded error may now masquerade undetected as a false-positive error; in essence, a situation similar to the so called "shadow effect," wherein an insufficient number of loci are sampled in a genetic study and individuals with similar genotypes are treated as the same individual (Mills et al. 2000, Waits and Leberg 2000).

Researchers are continually looking for new ways to mark individuals that facilitate long-term identification without recapture, while minimizing band loss, fading, and resighting errors. For larger-bodied shorebirds and waterbirds, one solution to the problems associated with band loss and fading has been the use of durable field-readable alphanumeric bands (Ottaway et al. 1984). However, for smaller-bodied birds, such alphanumeric bands are difficult to read and may be more prone to marker loss than non-numeric color bands. A possible alternative is the application of flags (i.e., plastic bands with a protruding rectangular tab of plastic) embossed with field-readable alphanumeric codes. For smaller-bodied birds, the greater surface area available on a flag allows larger alphanumeric characters and thus more readable codes than is possible with bands. For shorebirds, for which the international banding standard recommends the application of a colored flag as a regional marker, use of alphanumeric flags might be particularly attractive.

Piping Plovers (*Charadrius melodus*) are one example of a small-bodied migratory shorebird. Endemic to North America, these plovers are subject to extensive mark-recapture research and conservation management due to their federally threatened status. Since 2005, various entities conducting mark-recapture work on Piping Plovers have used colored plastic flags to indicate marking origin. We investigated the field-readability of a sample of nesting adult Piping Plovers each double-marked with a unique combination of colored leg bands and a single yellow alphanumeric flag embossed with a three-character code. Our ultimate objective was to assess whether or not the field-readable alphanumeric flags were as identifiable as unique color-band combinations. For each recorded encounter with a doubly-marked bird, we estimated the probabilities that observers would be able to read and record a color-band combination, an alphanumeric code, or both. For each reported marker, we determined whether the color-band combination or alphanumeric code was potentially valid (i.e., corresponded to a previously deployed marker). Finally, if both a color-band combination and alphanumeric code were read and were potentially valid, we assessed the frequency that both codes matched our banding records (i.e., represented the same bird).

METHODS

Study area. We conducted our study on riverine and reservoir habitats of the Missouri River extending southward from Lake Sakakawea in North Dakota to the headwaters of Lake Oahe, ~10 km south of Bismarck, North Dakota. Habitat on Lake Sakakawea consisted of irregular and dissected island or beach shorelines with numerous substrates, slopes, and aspects (Anteau et al. 2012). Habitat for Piping Plovers along the Garrison reach of the Missouri River (extending south from the Garrison Dam) occurred primarily on mid-channel low- to mid-elevation sandbars with some established woody vegetation.

Marking and resighting. We visited all sandbars, islands, and shoreline habitat every 2 to 3 d from mid-April through 31 July 2013 to locate nests and mark and resight Piping Plovers (see Shaffer et al. [2013] for more detailed information about field methods and the study area). We trapped adult plovers on nests during incubation using either a modified remote-controlled walk-in trap or bow-net. We banded Piping Plovers with four plastic color bands, a U.S. Geological Survey aluminum metal band, and a yellow plastic flag with a black three-character alphanumeric code (INTERREX, Lodz, Poland; http://interrex.nazwa.pl/colour-rings/info). In 2013, alphanumeric flag codes consisted of three-character combinations of the numerals 1-9 and letters A-E, H, L, N, and P. We placed two color bands on each leg below the tibiotarsal joint, the USGS metal band above the tibiotarsal joint of one leg, and the alphanumeric flag above the tibiotarsal joint of the other leg. Each plover could be individually identified using either its unique color-band combination or alphanumeric flag (Fig. 1).

Eight crews of 3–4 observers located marked birds using binoculars and spotting scopes. If observers were able to read the colors of all four color bands as well as the presence of a flag and metal band, the resighting event

Fig. 1. An example of a Piping Plover double-marked with a color-band combination and an alphanumeric flag with code B23. The color-band combination consists of an aluminum band above the tibiotarsal joint on the left leg followed by a cobalt-blue band over a black band located below the tibiotarsal joint on the same leg; the yellow alphanumeric flag is located above the tibiotarsal joint on the right leg with a black band over a yellow band located below the tibiotarsal joint on the same leg.

was coded as "complete color combo"; if the observer failed on any one of these accounts, the resighting event was coded as "incomplete color combo." If observers were able to read all three characters of the alphanumeric flag, they coded the resighting event as "complete ANF," otherwise they coded it as "incomplete ANF." When recording color-band combinations and codes of alphanumeric flags, observers wrote down the colors, characters, and positions of all markers seen, even if the combination or code was ultimately considered incomplete.

Estimating the probability of misidentification. We built multi-state models (Hestbeck et al. 1991, Brownie et al. 1993) in program MARK (White and Burnham 1999) to estimate the probability that 1) observers recorded a complete color-band combination or alphanumeric code for a double-marked bird, 2) the color-band combination or alphanumeric code read was one that had been deployed (i.e., valid, albeit not necessarily correct), and 3) if both a color-band combination and alphanumeric code were read, that they matched (i.e., were valid and had been deployed on the same bird). Multi-state models are typically used to simultaneously estimate the probability of survival (S), detection (p), and movement (ψ) of organisms in two or more states, which are traditionally either physical locations (e.g., islands or study areas) or states-of-being (e.g., breeder vs. non-breeder).

For each resighting event ($N = 671$) of a double-marked plover, we created a capture history consisting of four occasions with four states that represented the manner in which a double-marked plover could be resighted. Thus, in the context of this analysis, a "recapture occasion" actually represents different stages in the process of identifying a single plover during a single resighting event (Fig. 2); although each resighting event had four "occasions" (hereafter called "stages"), there could have been multiple resighting events for each individual plover. The four states included in our analysis consisted of: B – both alphanumeric flag and color combination, C – color combination only, F – alphanumeric flag only, and N – neither alphanumeric flag nor color combination. Although multi-state models allow simultaneously estimates of survival, detection, and movement at each stage of the resighting process, we defined models so that all survival and detection probabilities could be fixed to 1.

All capture histories began with the observation of a bird wearing both a color-band combination and an alphanumeric flag and were coded as B on occasion 1 (Fig. 2, stage 1). From this stage, a transition (ψ) to one of four states was possible for occasion 2, depending on whether an observer had read only the complete color-band combination (state C, ψ_{BC1}), only the complete alphanumeric flag (state F, ψ_{BF1}), both the complete color-combination and alphanumeric flag (state B, ψ_{BB1}), or neither (state N, ψ_{BN1}) (Fig. 2, stage 2). Observations that resulted in neither marker being read were censored from further stages of analysis (i.e., their capture histories were coded as BN.).

If only one of the two markers on a double-marked plover was successfully read, there were only two possible transitions to stage 3, i.e., the complete combo/code read could either be valid (ψ_{CC2} or ψ_{FF2}) or invalid (ψ_{CN2} or ψ_{FN2}). We defined color-band combinations and alphanumeric codes read by observers as valid if they had actually been deployed, or invalid if they had not been deployed. If only one marker was

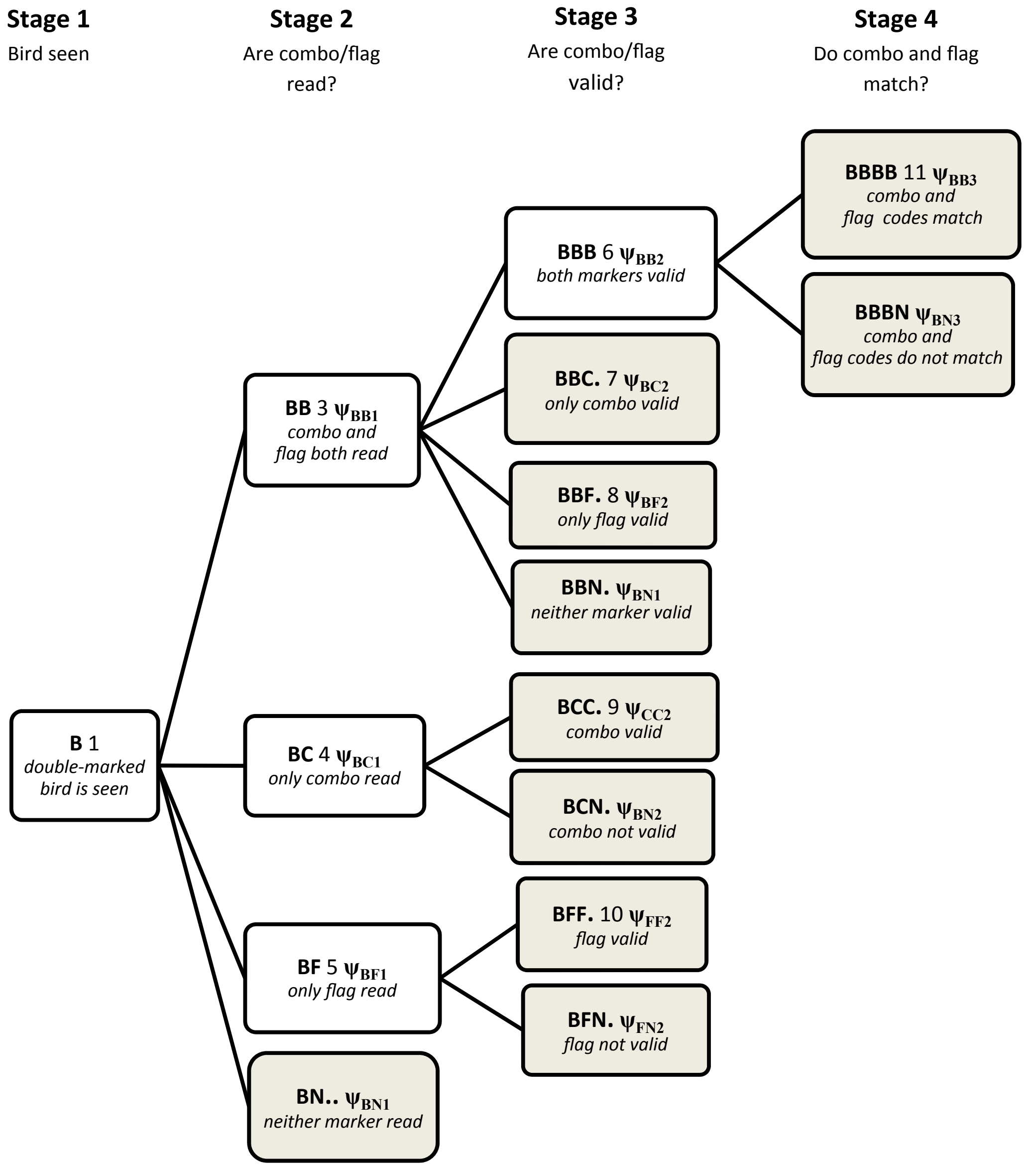

Fig. 2. Multi-state model for the process of resighting a double-marked Piping Plover consisted of four successive stages. The states are: B = both color combo (combo) and alphanumeric flag (flag) code, C = combo only, F = flag only, and N = neither. The model consists of four "occasions" that refer to sequential stages of the resighting process, with Stage 1 conditional on observation of a bird with both combo and flag (conditional probability = 1; birds not known to have both marker types are excluded). Stage 2 models whether the observer can successfully record (read) a full description of the combo and/or flag. Stage 3 models whether this description is potentially valid (i.e., matches the record of a previously banded bird, but it might not be the correct bird). Finally, Stage 4 models whether the recorded combo and flag match the banding records for the same bird (thus providing near certainty that the data are correct). At each node, we provide the cumulative capture history, the parameter index number (see Appendix), and the likelihood of arriving at that node from the previous node (e.g., the probability of arriving at BBBB is $1\times \psi_{BB1} \times \psi_{BB2} \times \psi_{BB3}$). Terminal nodes are shaded and for sequences that end before Stage 4, the capture history ends in periods to denote right censoring. The probabilities of all terminal nodes sum to 1, as do the probabilities of all branches originating from a single node. Detection failure (N) was always estimated via subtraction.

read, then transition probabilities to the other marker type or to both markers were fixed to 0, and the final capture occasion was censored (i.e., "." was entered for the final occasion; Fig. 2).

If both markers were read at stage 2, then a transition to any one of four outcomes was possible at stage 3, depending on whether only the color combination was valid (ψ_{BC2}), only the alphanumeric flag code was valid (ψ_{BF2}), both the color combination and alphanumeric flag code were valid (ψ_{BB2}), or neither was valid (ψ_{BN2}) (Fig. 2, stage 3). If the color combination and alphanumeric flag code were both valid at stage 3 (i.e., state B), then there were two possible outcomes in occasion 4: the combo and code both belonged to the same banding record and therefore "matched" (ψ_{BB3}), or the combo and code did not match (ψ_{BN3}) and therefore could not represent the same bird.

We tested four hypotheses about similarity of errors between color bands and alphanumeric flags: 1) the probability of successfully reading a marker combination did not differ for alphanumeric flags versus color bands, and 2–4) the probability of reading a valid combination did not differ between alphanumeric flags versus color bands when 2) only one marker type was read, 3) both marker types were read, or 4) regardless of whether one or both markers were read. Because only one of these hypotheses was directly testable in MARK using alternate models, we used Z-tests to contrast derived parameters with standard errors estimated using the delta method (Powell 2007). Values are presented as means ± SE.

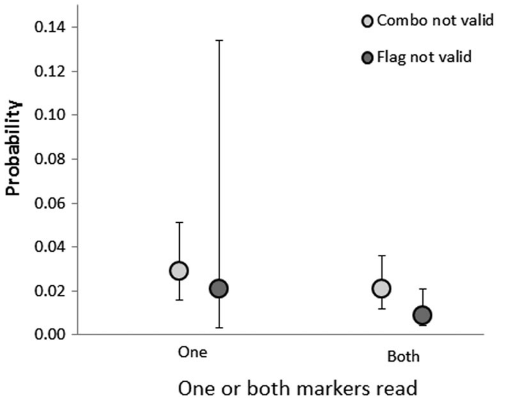

Fig. 3. The probability that a color combination (Combo) or alphanumeric flag (Flag) code was read incorrectly, depending on if it was the only marker read on a bird (One) or both markers were read (Both).

RESULTS

In 2013, 309 adult Piping Plovers were double-marked with an alphanumeric flag and a color-band combination and subsequently resighted over the course of 1019 separate resighting events. Observers read both a complete color-band combination and alphanumeric code more than half the time (56.7%), but, when observers read only one marker, they were more likely to read a color-band combination (37.4%) than an alphanumeric flag code (4.7%). The combined probability of reading a color-band combination ($\psi_{BB1} + \psi_{BC1}$) was 0.941 ± 0.007 versus 0.614 ± 0.015 for the combined probability ($\psi_{BB1} + \psi_{BF1}$) of reading an alphanumeric flag ($Z = 19.3$, $P < 0.0001$).

If observers were only able to read one of the two markers, they failed to read a valid color-band combination 2.9 ± 0.9% of the time, versus 2.1 ± 2.1% of the time for alphanumeric flags ($Z = 0.4$, $P = 0.72$; Fig. 3, Table 1). When observers read both marker types, they usually (97.0%) read valid combinations and codes (Table 1, ψ_{BB_2}). However, if a mistake occurred, it was twice as likely to involve a color-band combination (2.1%) as an alphanumeric flag (0.9%; Table 1), although the difference was not significant ($Z = 1.7$, $P = 0.087$). Error rates were similar when only one versus both markers were read, so we estimated the combined probability of recording an invalid color-band combination (2.4 ± 0.7%) versus an invalid flag (1.0 ± 0.5%), a difference that was not significant ($Z = 1.7$, $P = 0.097$). There were no cases where both the alphanumeric and color-band combinations were invalid. When observers were able to read both a complete color-band combination and an alphanumeric flag, and both were valid, they matched 95.5% of the time (Table 1, ψ_{BB_3}).

DISCUSSION

In mark-recapture studies that depend on genetic markers or photo identification, misidentification errors go undetected, adding apparently

Table 1. Probabilities of reading a color-band combination (combo) or alphanumeric flag (flag) on a double-marked Piping Plover. Read means a full description was recorded. Valid means the description matches a previously deployed marker. Match means the combo and flag were both consistent with the same previously marked bird. Estimates were generated using the fully parameterized model as described in Supplemental Appendix S1.

Parameter	description	*N*	Est.	LCI	UCI
Stage 1	Double-marked bird is observed		1.000	-	-
Stage 2	Are combo/flag read?				
ψ_{BB_1}	Both read	578	0.567	0.537	0.597
ψ_{BC_1}	Only combo read	381	0.374	0.345	0.404
ψ_{BF_1}	Only flag read	48	0.047	0.036	0.062
ψ_{BN_1}	Neither read	12	0.012	0.007	0.021
Stage 3	Are combo/flag valid?				
ψ_{CC_2}	Only combo read, combo valid	370	0.971	0.949	0.984
ψ_{CN_2}	Only combo read, not valid	11	0.029	0.016	0.051
ψ_{FF_2}	Only flag read, flag valid	47	0.979	0.866	0.997
ψ_{FN_2}	Only flag read, not valid	1	0.021	0.003	0.134
ψ_{BB_2}	Both read, both valid	561	0.970	0.953	0.982
ψ_{BC_2}	Both read, flag not valid	5	0.009	0.004	0.021
ψ_{BF_2}	Both read, combo not valid	12	0.021	0.012	0.036
ψ_{BN_2}	Both read, neither valid	0	0.000	0.000	0.000
Stage 4	Do combo and flag match?				
ψ_{BB_3}	Both read, both valid, both match	536	0.955	0.935	0.970
ψ_{BN_3}	Both read, both valid, but no match	25	0.045	0.030	0.065

new individuals to the study and resulting in inflated population-size estimates and deflated survival estimates. However, for mark-recapture studies where a known number of readable markers (e.g., tags and bands) are deployed, investigators do not have to contend with undetected misidentification errors in the same manner; reports of marker codes that were never deployed can easily be identified and excluded from analyses. If misreads result in a marker that was already deployed, these errors (i.e., false-positive errors) would likely be interpreted as a valid detection of the wrong animal. Unlike previous band resighting studies where data-cleaning activities were restricted to resighting errors that resulted in a non-existent or impossible combination/code (e.g., Weiss et al. 1991), our use of two resightable markers allowed us to identify when resighting errors resulted in reports of combinations or codes that did exist, but were still incorrect.

We found that when observers read valid color-band combinations and alphanumeric codes for a double-marked Piping Plover, the two markers matched 95.5% of the time. For 4.5% of encounters where markers did not match, we could not determine if errors resulted from misreading the alphanumeric flag or the color-band combination. It is unlikely these errors were shared equally by both marker types, given our finding that recorded color-band combinations were more likely to be invalid than recordings of alphanumeric flag codes. If errors were shared equally between marker types, then marker-specific error rates would be $1 - \sqrt{0.955}$, or 0.023. Conversely, if the error rate for each marker type was proportional to the rate at which non-existent combinations were observed (23 of 959 color-band observations, or 2.4% vs. 6 of 626 flags observations, or 1.0%), then we would have expected 2.4 times as many errors for color bands (C) as for flags (F), or $(1\text{-}C)^*(1\text{-}F) = 0.955$, with $C = 2.4^*F$. Solving these two joint equations gives $F = 0.013$ and $C = 0.032$. Given a sample of 309 marked birds, we can safely ignore the $(1/309)^2 = 0.0000105$ possibility that both markers were wrong, but fortuitously matched a previously marked plover.

The total resighting error in our study was about 7.5%, representing the sum of the three principal error types, i.e., misread color-band combination (2.1%), misread alphanumeric flag (0.9%), and mismatch between band and flag (4.5%). This error rate was comparable to the resighting errors reported in previous studies (e.g., 5%, Milligan et al. 2003; 6.3%, Weiss et al. 1991; 10%, Schwarz and Stobo 1999; 3.5–13.4%, Lavers and Jones 2008), although our frequency of detectable false-negative errors for either alphanumeric codes (0.9%) or color band combinations (2.1%) was far lower than reported previously. For example, Weiss et al. (1991) reported that the probability of misreporting neck-banded Canada Geese (*Branta canadensis*) was 6.3%. However, Weiss et al. (1991) were limited to detecting resighting errors of non-existent combinations, e.g., neck-band codes "resighted" when the neck band was known to have already been harvested or never deployed. As the authors noted, there was no way to discriminate between neck-band misreads that resulted in a valid neck-band code for a different goose (Weiss et al. 1991). If the results of our study are any indication, the frequency of total error could have been up to twice as high and the frequency of such undetected errors would increase with the number of valid neck-band codes.

Investigators who have measured error without discriminating among "types" of error, typically report error rates comparable to the total error rate we present here. Milligan et al. (2003) reported that resighting errors for color-banded passerines were highly dependent on observer ability, ranging from 7-54%. Because this was a controlled study where the true color band combinations were known, these rates reflect total error. Similarly, in a study of Razorbills (*Alca torda*) where field-readable leg-band codes were known, Lavers and Jones (2008) reported error rates ranging from 0.035 to 0.134, depending on observer experience and conditions.

False-positive errors seem to become a more pernicious issue as the number of valid combinations available and observer familiarity with a system increase. For example, Lavers and Jones (2008) found that, because 12,000 Razorbills had been banded from a continuous string since 1980, 94% of the resighting errors made corresponded to a valid band number. Similarly, Schwarz and Stobo (1999) reported a misreading rate of 0.10 for branded seals, with nearly all errors resulting in a valid brand number (i.e., a false-positive error), potentially because observers were aware of all possible combinations and thus were unlikely to report a number they knew did not exist. This rationale may also explain why the probability of false-negative errors was so low in our study, whereas the number of false-positive errors (Table 1, $\psi_{BN_3} = 0.045$) was much larger.

We believe the results we present here represent one of the more conservative measures of color-band resighting error. All markers used in our study were newly applied in 2013, observers were experienced and knowledgeable of the banding scheme, band loss was non-existent (i.e., all band combinations were complete), color fading was negligible, and the number of valid band combinations or codes available was limited. In short, both marker types were at their most readable. Over time, we suspect that resighting errors would increase, and disproportionately so for color bands, for a number of reasons. As additional markers are deployed in future years, the increased number of deployed band combinations would increase the potential for false-positive error. Fading and loss are likely to be more problematic for color bands than for alphanumeric flags, leading to increasing difficulty with accurately reading color bands. Resighting error rate is unlikely to be constant throughout the duration of a band-resighting program, particularly as new bands are deployed and the condition of previously deployed bands deteriorates. Thus, incorporating measures for quantifying error rate may improve accuracy and precision of vital rates estimated through band resightings. Alphanumeric flags can help meet this need in programs such as ours, with the additional benefit of continuing to allow individual identification after color bands deteriorate or in the face of color blindness (Bear et al. 1989), even if recapture and physically double-checking the metal band becomes impossible.

The use of a double marker has long been extolled for bird mark-recapture studies, typically in the form of a "permanent" metal band (Kendall et al. 2009), and here we simply extended this concept to include a second field-readable marker to allow for resighting error correction without recapture. Field-readable alphanumeric flags offer a promising auxiliary marker for estimating and potentially adjusting for the frequency of false-positive resighting errors that otherwise can bias demographic estimates (Arnason and Mills 1981, Mills et al. 2000, Creel et al. 2003, Yoshizaki et al. 2009, Morrison et al. 2011). Lastly, a common concern in banding studies is the potential to exhaust the number of available field-readable combinations; two commonly employed techniques to deal with this are to either increase the number of color bands used in a banding scheme or to limit use of uniquely identifiable color-band combinations to breeding adults. The first alternative is never popular with permitting officials and, in addition, adding the number of color bands used in a combination can be problematic over time because investigators lose the ability to distinguish complete combinations from those that are missing bands. Limiting the application of unique field-readable markers to only breeding birds results in a loss of information on the demographics of non-breeding or pre-breeding individuals (Saunders et al. 2014). Field-readable alphanumeric codes could offer a solution to the "finite band combination" problem because with a single three-character alphanumeric flag 42,875 permutations of the numerals 1–9 and A–Z are possible (i.e., the 35 characters can occur in three positions, $35^3 = 42,875$).

ACKNOWLEDGMENTS

This study was funded by the U.S. Army Corps of Engineers' Missouri River Recovery Program through financial and logistical support from the Corps' Omaha District Threatened and Endangered Species Section and Garrison Project Office. We are grateful for technical support by the USGS Northern Prairie Wildlife Research Center Missouri River Least Tern and Piping Plover Research Team. We thank B. Hill, M. Ring, D. Toy, C. Huber, C. Kruse, and C. Aron for help with project planning, logistics, or data collection and analysis. We also thank many field technicians and crew leaders for their assistance with data collection. Lastly, we are indebted to an anonymous reviewer as well as D. Johnson, T. Shaffer, and M. Sherfy for comments that improved the manuscript. Our field protocols were approved by the USGS Northern Prairie Wildlife Research Center Animal Care and Use Committee. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

LITERATURE CITED

Anderson, D. R., K. P. Burnham, and G. C. White. 1985. Problems in estimating age-specific survival rates from recovery data of birds ringed as young. Journal of Animal Ecology 54: 89–98.

Anteau, M. J., M. H. Sherfy, and M. T. Wiltermuth. 2012. Selection indicates preference in diverse habitats: a ground-nesting bird (*Charadrius melodus*) using reservoir shoreline. PLoS ONE 7: e30347.

Arnason, A. N., and K. H. Mills. 1981. Bias and loss of precision due to tag loss in Jolly-Seber estimates for mark-recapture experiments. Canadian Journal of Fisheries and Aquatic Science 38: 1077–1095.

Bear, G. D., G. C. White, L. H. Carpenter, R. B. Gill, and D. J. Essex. 1989. Evaluation of aerial mark-resighting estimates of elk populations. Journal of Wildlife Management 53: 908–915.

Brownie, C., J. E. Hines, J. D. Nichols, K. H. Pollock, and J. B. Hestbeck. 1993. Capture-recapture studies for multiple strata including non-Markovian transitions. Biometrics 49: 1173–1187.

Conn, P. B., W. L. Kendall, and M. D. Samuel. 2004. A general model for the analysis of mark-resight, mark-recapture, and band-recovery data under tag loss. Biometrics 60: 900–909.

Creel, S., G. Spong, J. L. Sands, J. Rotella, J. Zeigle, L. Joe, K. M. Murphy, and D. Smith. 2003. Population size estimation in Yellowstone wolves with error-prone noninvasive microsatellite genotypes. Molecular Ecology 12: 2003–2009.

Diefenbach, D. R., and G. L. Alt. 1998. Modeling and evaluation of ear tag loss in black bears. Journal of Wildlife Management 62: 1292–1300.

Hastings, K. K., L. A. Hiby, and R. J. Small. 2008. Evaluation of a computer-assisted photograph-matching system to monitor naturally marked harbor seals at Tugidak Island, Alaska. Journal of Mammalogy 89: 1201–1211.

Hestbeck, J. B., J. D. Nichols, and R. A. Malecki. 1991. Estimates of movement and site fidelity using mark-resight data of wintering Canada Geese. Ecology 72: 523–533.

Juillet, C., R. Choquet, G. Gauthier, and R. Pradel. 2010. A capture-recapture model with double-marking, live and dead encounters, and heterogeneity of reporting due to auxiliary mark loss. Journal of Agricultural, Biological, and Environmental Statistics 16: 88–104.

Kendall, W. L., S. J. Converse, P. F. Doherty, Jr., M. B. Naughton, A. Anders, J. E. Hines, and E. Flint. 2009. Sampling design considerations for demographic studies: a case of colonial seabirds. Ecological Applications 19: 55–68.

Lavers, J. L., and I. L. Jones. 2008. Assessing the type and frequency of band resighting errors for Razorbill *Alca torda* with implications for other wildlife studies. Marine Ornithology 36: 19–23.

Lebreton, J. D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecological Monographs 62: 67–118.

Lincoln, F. C. 1930. Calculating waterfowl abundance on the basis of banding returns. USDA Circular 118: 1–4.

Link, W. A., J. Yoshizaki, L. L. Bailey, and K. H. Pollock. 2010. Uncovering a latent multinomial: analysis of mark-recapture data with misidentification. Biometrics 66: 178–185.

Lukacs, P. M., and K. P. Burnham. 2005a. Research notes: estimating population size from DNA-based closed capture-recapture data incorporating genotyping error. Journal of Wildlife Management 69: 396–403.

———, and ———. 2005b. Review of capture-recapture methods applicable to noninvasive genetic sampling. Molecular Ecology 14: 3909–3919.

Marshall, A. D., and S. J. Pierce. 2012. The use and abuse of photographic identification in sharks and rays. Journal of Fish Biology 80: 1361–1379.

Miller, D. A., J. D. Nichols, B. T. McClintock, E. H. Campbell Grant, L. L. Bailey, and L. A. Weir. 2011. Improving occupancy estimation when two types of observational error occur: non-detection and species misidentification. Ecology 92: 1422–1428.

Milligan, J. L., A. K. Davis, and S. M. Altizer. 2003. Errors associated with using colored leg bands to identify wild birds. Journal of Field Ornithology 74: 111–118.

Mills, L. S., J. J. Citta, K. P. Lair, M. K. Schwartz, and D. A. Tallmon. 2000. Estimating animal abundance using noninvasive DNA sampling: promise and pitfalls. Ecological Applications 10: 283–294.

Morrison, T. A., J. Yoshizaki, J. D. Nichols, and D. T. Bolger. 2011. Estimating survival in photographic capture-recapture studies: overcoming misidentification error. Methods in Ecology and Evolution 2: 454–463.

Nichols, J. D., and J. E. Hines. 1993. Survival rate estimation in the presence of tag loss using joint analysis of capture-recapture and resighting data. In: Marked individuals in the study of bird populations (J. D. Lebreton and Ph. M. North, eds.), pp. 141-155. Birkhauser, Basel, Switzerland.

Ottaway, J. R., R. Carrick, and M. D. Murray. 1984. Evaluation of leg bands for visual identification of free-living Silver Gulls. Journal of Field Ornithology 55: 287–308.

Petersen, C. G. T. 1896. The yearly immigration of young plaice into the Limfjord from the German Sea. Report of the Danish Biological Station 6: 1–48.

Pollock, K. H., J. D. Nichols, C. Brownie, and J. E. Hines. 1990. Statistical inference for capture-recapture experiments. Wildlife Monographs 107: 3–97.

Powell, L. A. 2007. Approximating variance of demographic parameters using the delta method: a reference for avian biologists. Condor 109: 949–954.

Royle, J. A., and W. A. Link. 2006. Generalized site occupancy models allowing for false positive and false negative errors. Ecology 87: 835–841.

Saunders, S., T. W. Arnold, E. A. Roche, and F. J. Cuthbert. 2014. Age-specific survival and recruitment of Piping Plovers (*Charadrius melodus*) in the Great Lakes region. Journal of Avian Biology 45, in press. DOI: 10.1111/jav.00319.

Schwarz, C. J., and W. T. Stobo. 1999. Estimation and effects of tag-misread rates in capture-recapture studies. Canadian Journal of Fisheries and Aquatic Science 56: 551–559.

Shaffer, T. L., M. H. Sherfy, M. J. Anteau, J. H. Stucker, M. A. Sovada, E. A. Roche, M. T. Wiltermuth, T. K. Buhl, and C. M. Dovichin [online]. 2013. Accuracy of the Missouri River Least Tern and Piping Plover monitoring program—considerations for the future. U.S. Geological Survey Open-File Report 2013-1176. http://pubs.usgs.gov/of/2013/1176/ (Accessed 14 April 2014).

Spendelow, J. A., J. Burger, I. C. T. Nisbet, J. D. Nichols, J. E. Hines, H. Hays, G. D. Cormons, and M. Gochfeld. 1994. Sources of variation in loss rates of color bands applied to adult Roseate Terns (*Sterna dougallii*) in the Western North Atlantic. Auk 111: 881–887.

Stevick, P. T., P. J. Palsboll, T. D. Smith, M. V. Bravington, and P. S. Hammond. 2001. Errors in identification using natural markings: rates, sources, and effects on capture-recapture estimates of abundance. Canadian Journal of Fisheries and Aquatic Science 58: 1861–1870.

Waits, J. L., and P. L. Leberg. 2000. Biases associated with population estimation using molecular tagging. Animal Conservation 3:191–199.

Weiss, N. T., M. D. Samuel, D. H. Rusch, and F. D. Caswell. 1991. Effects of resighting errors on capture-resight estimates for neck-banded Canada Geese. Journal of Field Ornithology 62: 464–473.

White, G. C., and K. P. Burnham. 1999. Program MARK: Survival estimation from populations of marked animals. Bird Study 46: S120–S139.

Wright, J. A., R. J. Barker, M. R. Schofield, A. C. Frantz, A. E. Byrom, and D. M. Gleeson. 2009. Incorporating genotype uncertainty into mark-recapture-type models for estimating abundance using DNA samples. Biometrics 65: 833–840.

Yoshizaki, J., K. H. Pollock, C. Brownie, and R. A. Webster. 2009. Modeling misidentification errors in capture-recapture studies using photographic identification of evolving marks. Ecology 90: 3–9.

Zimmerman, G. S., W. L. Kendall, T. J. Moser, G. C. White, and P. F. Doherty, Jr. 2009. Temporal patterns of apparent leg band retention in North American geese. Journal of Wildlife Management 73: 82–88.