Report on the 2018 Aleutian Tern Conservation Planning Meeting

May 31, 2018

Prepared for:
Aleutian Tern Technical Committee
Sponsored by:
National Fish & Wildlife Foundation

Prepared by:
Trent L. McDonald and **Jason D. Carlisle**
Western EcoSystems Technology, Inc.
415 West 17th Street, Suite 200
Cheyenne, Wyoming 82001

Executive Summary

The Aleutian Tern Technical Committee convened an Aleutian Tern Conservation Planning Meeting in Anchorage, Alaska, January 26-27, 2018. The meeting's overarching goal was to work towards methods for assessing statewide Aleutian Tern population size and trend. The primary question addressed at the meeting was whether and how within-year population abundance at the local (i.e., colony) scale could be estimated. Following two days of presentations and discussion, workshop participants recommended pilot testing three direct and one indirect method for estimating abundance at a few non-randomly selected colonies during summer 2018. Meeting participants recommended testing the following direct count methods: (1) optical ground-based counts of flying birds; (2) counts of flying birds in ground-based photos; and (3) counts of birds on the ground in low-altitude (10-20 meters) photography obtained by unmanned aerial vehicles. The recommended indirect method will attempt to correlate the number and seasonal timing of recorded Aleutian Tern calls with abundance or nest density at select colonies. The main challenges facing the overall monitoring program include how to define a “colony,” which metric to target (abundance of nests, breeding pairs, or adults), estimation of species ratios at mixed-species colonies, the seasonal timing of surveys, the number of birds off-colony during the survey, and the proportion of birds flushing during counts of in-flight individuals. Once some of the questions surrounding estimation at a single colony are resolved, the Technical Committee will focus on and develop spatial and temporal sampling methods necessary to expand estimates to the entire state. This report lays out the general protocol for the 2018 pilot methods and lists questions to be considered. We also give initial thoughts on how the monitoring program could be expanded to the entire state. The motivating framework, agenda, list of attendees, and detailed notes taken at the meeting are included as appendices.

Table of Contents

1. Introduction

The Aleutian Tern (ALTE, *Onychoprion aleuticus*) is an uncommon seabird that nests in coastal areas of Alaska and Russia (North 2013). A recent analysis of past abundance data estimated that the number of ALTE at known colonies in Alaska has declined 8.1% annually (95% credible interval = 10.7 - 5.5%) since 1960 (Renner et al. 2015). The drivers of this population decline are unclear, but may include partial habitat modification, predation, egg harvesting, pesticides or other contaminants, changes in marine prey-base, human disturbance at nest sites, and/or overwinter mortality (Aleutian Tern Conservation Working Group [ATCWG] 2007 and ATCWG 2015, BirdLife International 2018, North 2013). In light of the rapidity of perceived population declines, the status of ALTE was recently changed to "vulnerable" by the International Union for Conservation of Nature and Natural Resources (BirdLife International 2018).

Concern over ALTE status has led to a mobilization of conservation partners in recent years. As part of this mobilization, a group of professional seabird researchers and managers met informally in 2007 (ATCWG 2007). Based on collaborative work after that meeting, the group released an updated summary of information needs and research priorities in 2015 (ATCWG 2015). The following year, the Pacific Seabird Group formed the Aleutian Tern Technical Committee ("Technical Committee") to begin addressing information needs. After it became apparent that a larger multi-day workshop was required to discuss issues, the Technical Committee established a meeting framework (Appendix A) and convened the first ALTE Conservation Planning Meeting in Anchorage, Alaska on January 26-27, 2018 (agenda attached as Appendix B). The meeting's participants (Appendix C) represented state agencies, federal agencies, academic institutions, non-profit foundations, consulting firms, and private individuals. The meeting's overarching goal, objective, and focal question were as follows:

Overarching Goal: To design a statewide monitoring method that generates an unbiased estimate of ALTE abundance in Alaska (see also Appendix D).

Objective: To facilitate a common understanding of alternative sampling and population estimators for ALTE, including assumptions, advantages, and limitations, ultimately leading to recommended methodologies for future population monitoring.

Focal Question: The primary question addressed at the meeting and in this report was, "How do we best estimate within-year population abundance of ALTE at the local (i.e., colony) scale?"

This is a report on the outcomes, considerations, and final recommendations made during the planning meeting in January 2018. The next section, entitled **Outcomes**, contains a description of the meeting's general actionable items. The second section, **Site Design Considerations**, discusses considerations that need to be addressed by all pilot methods. The following section, **Recommended Pilot Methods in 2018**, contains specific recommendations for pilot testing of field methods in 2018. Finally, **Preliminary Recommendations for the Spatial and Temporal Designs** contains recommendations for statewide surveys to be refined after pilot testing of methods is complete.

2. Outcomes

During two days of presentations and discussion, meeting participants recognized that many logistical and feasibility questions surround the *site design* component of the overall monitoring plan. The site design defines *what* is measured at a colony and *how*. Participants agreed that an estimate of total colony abundance was the most desirable estimation goal of the site design (i.e., *what* = abundance), but significant questions remained regarding whether this goal could actually be achieved. If adequate estimates of total abundance are not achievable, other reasonable estimation goals include the number of breeding pairs, number of nesting birds, and other metrics. *How* to reliably estimate total abundance at a colony was the key question and no consensus method emerged. Going forward, the *spatial and temporal* components of the overall monitoring plan (McDonald, 2015) are difficult to consider without a firm site design in place. Consequently, participants recommended the following outcomes:

1. **Conduct pilot testing in 2018.** Due to uncertainties of the site design, workshop participants recommended pilot testing field methods that target estimation of ALTE abundance at a single colony. During summer 2018, three direct and one indirect methodology will be pilot tested at selected colonies/sites. The goal of pilot testing will be evaluation of each method's feasibility and cost.

2. **Evaluate piloted methods**. Following summer 2018 pilot testing, the effectiveness and cost of each method for generating unbiased colony abundance estimates will be evaluated.

3. **Reconvene and determine next steps**. Once the final report on summer 2018 activities is complete, the Technical Committee will convene a meeting to evaluation the piloted methods and work toward determining method(s) for future population monitoring.

3. Site Design Considerations

Meeting participants discussed many site design considerations and fieldwork options that bear on the overall monitoring method. Some considerations largely impact one method (e.g., overcounting is primarily a concern for direct counts), while others impact all methods. In this section, we list and briefly discuss survey design considerations that need to be addressed by all pilot methods proposed for summer 2018.

3.1. Can we accurately assess the Aleutian Tern / Arctic Tern species ratio at mixed species colonies?

Mixed species colonies contain both ALTE and Arctic Terns (ARTE, *Sterna paradisaea*). All piloted field methods will need to differentiate ALTE and ARTE or estimate a species ratio at mixed species colonies. Development of a species ratio, either at the time of survey or over a longer time frame, must rely on a repeatable quantitative method that can differentiate ALTE and ARTE. For example, a species ratio might be developed from certain flush counts, assuming ALTE and ARTE can be differentiated visually or aurally during flight. Or, past efforts indicate that ALTE and ARTE nests can be differentiated visually while walking transects and this information can be used to develop a species ratio assuming the ratio of ALTE to ARTE nests is the same as the ratio of ALTE to ARTE adults.

3.2. What constitutes a “colony”?

The definition of a colony has been one of the most vexing problems facing the ALTE monitoring program. Several colony definitions have been proposed and used in the past (e.g., North 2013, Pyare et al. 2013). For purposes of obtaining an accurate regional abundance estimate, we recommend differentiation between a *statistical colony* and a *biological colony.* A *statistical colony* can and should be defined by practical sampling considerations and used as the sample unit for ALTE abundance estimation. The statistical colony may not agree with a biologist's view of the area that functions as a *biological colony*. That is, it is possible and reasonable to construct different definitions for the *statistical colony* and the *biological colony* at a site used by ALTE. For analytical purposes, the unit of replication is the statistical colony. In the remainder of this report, we make a distinction between statistical colonies and biological colonies. The statistical colony definition need only be workable over a long time frame, easy to access during field visits, and easy to count. In contrast, the biological definition of a particular colony can change through time, and biological colonies can die or move during the course of study. It will be extremely useful to use the same statistical colony definition at all biological colonies. With these thoughts in mind, we list and generally discuss three potential working definitions of statistical colonies. We discuss more specific considerations regarding these definitions as part of each protocol in Section 4 below.

A. *Ground-based point and distance definition*: Under this definition, the statistical colony is associated with a point, likely near the center of a biological colony. Placement of the point in space is arbitrary and should be established by biologists familiar with the site. Once a center point is established, it should be permanent. The count method (either flush counts or aerial photos from drones) would use the center point as its reference or origin. For example, observers could attempt to flush all birds within X meters of the point or the drone could survey a $X \times X$ meter quadrat surrounding the point.

B. *Interstitial distance definition:* It may be possible to define a statistical colony as birds that are within X meters of each other, either in flight or on the ground. For example, a drone could continue surveys until a gap of X meters between individuals is observed. This definition may be difficult to implement because it must be possible for observers to measure X between individuals, in all directions around the colony, during sampling. The geographic size of the statistical colony under this definition would vary with each field visit.

C. *Landmark-based definition*: Assuming suitable vantage points can be identified, it may be possible to define a statistical colony using permanent landmarks. For counts of flying birds, landmarks must be within an observer's background field of view. For ground counts, landmarks can be nearly any permanent ground feature (e.g., river, cliff, beach). Boundaries of the colony should be large enough to accommodate among-year shifts in distribution. Visibility of all birds within the boundaries should be similar. Under this definition, the geographic shape of statistical colonies would vary among biological colonies, but the geographic size of statistical colonies would be constant across visits to the same biological colony.

3.3. What happens when the statistical colony definition stops working?

In general, it is a bad idea to change statistical colony definitions (sample units) during monitoring. Nonetheless, changes from one statistical colony definition (e.g., a ground-based point) to another (e.g., landmark-based) at a biological colony may be necessary during a long-term study. When the statistical colony definition no longer works, there are two cases to consider:

A. No birds: The program may have multi-season and multi-year evidence that a colony is no longer functional. Birds may not visit the site, habitat at the site may have changed, or use may have shifted beyond the colony's established boundaries. When this happens to a statistical colony that was previously counted, we recommend continuing to apply the protocol (i.e., count) birds on the old statistical colony until the sampling frame is updated (see Section 5.1). The assumedly low or zero count on the old statistical colony should still be entered into the database until the sampling frame is updated. If no birds are sighted across several field visits and across several seasons, further field visits may not be necessary. When the sample frame is updated, the old statistical colony can be dropped and a new statistical colony covering the new biological colony (if it exists) can be added to the frame and potentially sampled. Sampling would not continue at the biological colony unless one of the statistical colonies in that biological colony was chosen during sample selection.

B. Logistical changes: When it becomes necessary to change the statistical colony definition mid-study for logistical reasons (e.g., budget cuts, lost access to vantage points) and birds continue in both the old and new statistical definitions, we advocate stopping or "statistically killing" the old statistical colony and immediately starting counts on the new statistical colony. It is important that the old and new statistical colony ID's be separate and unambiguous so that analyses can take the change into account. When the sampling frame is updated (see Section 5.1), the ID of the old statistical colony can be removed.

3.4. How much disturbance do surveys cause?

All piloted methods need to consider the amount of disturbance they cause. Questions to ask and answer include, “Does sampling increase the probability of nest abandonment?” and “Does sampling and other activities at the colony provide cues to avian and mammalian predators that tern nests are available for depredation?” The amount of harm done to colonies, if any, by each method is largely unknown and should be quantitatively measured. Separate studies that closely monitor nest success and fledgling survival at colonies with and without anthropogenic disturbance are likely necessary to estimate these effects.

3.5. What proportion of the birds are at sea feeding when we visit?

ALTE generally forage at sea and return to the colony to socialize, incubate eggs, feed young and roost. When individuals are away from the colony, they are not available to be counted. None of the pilot field methods intrinsically adjust for less than 100% availability. Whatever colony sampling method is used, separate methods must estimate the proportion of colony inhabitants available during the visit. Nest camera photos, acoustic data, and telemetry studies can all inform availability patterns at a single colony and perhaps multiple colonies. Linda Welch, a workshop participant, reported successful use of small VHF tags (nanotags, from Lotek and others) and fixed receiver stations to estimate flux of individuals on colonies in Maine (Bluso-Demers 2010, Loring et al. 2017, Taylor, et al. 2017).

3.6. Are floaters adequately counted?

Non-breeding individuals (“floaters”) associated with a colony may not locally co-locate with breeding individuals, may have different attendance patterns, and may be more likely to switch colonies. All colony count methods need to consider the number and location of floaters, and in particular whether both loafing (if known) and nesting areas should be surveyed.

4. Recommended Pilot Methods in 2018

An outcome of the ALTE Conservation Planning Meeting was a recommendation to field test three direct and one indirect method for estimating single-colony abundance (breeding + non-breeding birds, adults + sub-adults). If an unbiased estimate of ALTE abundance can be determined at a properly drawn sample of statistical colonies (for “properly drawn sample,” see Section 5), the monitoring program can estimate both the total number of ALTE in Alaska and trends in statewide abundance. If an unbiased estimate of colony abundance *cannot* be determined, the program can still estimate trends in statewide ALTE abundance (provided a properly drawn sample of colonies in the state is consistently measured with the same protocol) but cannot estimate the total number of ALTE in the state. Meeting participants generally agreed that the ability to estimate both current status and current trends are high priorities, but that trends were both more important for management and perhaps more feasible in the field.

Given the goal of unbiased estimation of ALTE abundance at a single colony, participants agreed to field test the methods outlined in Sections 4.1 - 4.4 below.

4.1 Ground-based Direct Counts

The existing ALTE “statewide protocol” (Pyare et al. 2013) uses a ground-based quick-count protocol to estimate abundance at a colony. From independently replicated flush counts (counts of flying birds) collected in accordance with the “statewide protocol” it should be possible to apply the method of unreconciled double-observer counts (Royle 2004, Riddle et al. 2010), or a variant, to obtain a less-biased or unbiased estimate of total colony abundance. In general, the unreconciled double-observer count method assumes no double counting of individuals, so that the maximum among all independent counts is an accurate estimator of the abundance. This assumption is problematic when counting flushed terns and is potentially exacerbated in mixed-species flocks, but the extent of overcounting is unknown. Whatever the extent of overcounting, it should decrease with observer experience and additional independent counts. Furthermore, it may be possible to amend the hierarchical model of Riddle et al. (2010) in a way that allows for overcounting or for the fact that field counts are only approximate, and this should decrease bias in the overall method.

4.1.1 Direct-count Protocol

The following direct-count protocol closely follows the existing “statewide protocol” and is relatively general. This description is designed to provide a clear understanding of the method, but is not a rigorous field protocol by itself.

- Colonies should be visited at least once at a time when nesting activity is anticipated to be high. Nesting activity and colony residency varies from late April to mid-July.
- Visits can be conducted any time during daylight hours.
- Colonies with easy access (e.g., on Kodiak Island’s road system) can and should be visited multiple times in a single season. Multiple visits will provide information on nesting phenology and seasonal variation in colony attendance.
- At least three people should visit the colony at the same time (two observers and a flusher). Additional observers are recommended.
- If multiple counts are performed in a single day, flushes should be a minimum of two hours apart to reduce impacts and to let birds re-settle.
- If counting a new colony, observers should pre-scan the colony from a reasonable distance (50 meters) for 10 minutes to determine whether it is a mixed colony.

- At all mixed-species colonies, observers should pre-scan the colony from a reasonable distance (50 meters) for 10 minutes and record independent estimates of the proportion of ALTE at the colony.
- Following pre-scan, observers should position themselves at different vantage points around the colony up to 200 meters apart.
- When ready, the flusher should approach the colony and flush birds.
- All observers should take as long as necessary to count birds in flight. If hundreds of birds are in flight, calibrate a search image for “blocks” of 20 individuals. In this case, each observer should independently estimate the number of blocks of 20 individuals in flight. If the colony is estimated to have fewer than 100 birds, raw counts should be performed and recorded.
- All people should move away from the colony as quickly as possible after the flushing event and otherwise seek to minimize disturbance.

4.1.2 Direct-count Considerations

Ground-based flush-counts are quick, easy, and relatively low-tech methods for estimating colony abundance. The key to the method's success is accurate counting of flying birds (i.e., unbiased, meaning the average is correct over multiple applications). However, this is one of the protocol's most difficult components. In addition to general considerations listed in Section 3, we list here considerations specific to ground-based flush counts.

A. **Flush method and the statistical colony definition**. The method used to flush birds and the statistical colony definition are linked. Here are some potential flush methods for the three statistical colony definitions given in Section 3.2.

1. *Ground-based point and distance definition*: One method of flush-counting when the statistical colony is a defined distance from a geographic point is for the flusher to walk in a straight line to the center point and continue past it to a distance of X meters (*X* varies by colony). The same line through the center point must be walked every time the colony is counted. Under an alternative flush-count method, the flusher could spend T minutes (*T* varies by colony) circling the center point, starting at a distance of X meters, and spiraling inward. Variable colonies shapes and extents (e.g., due to beach, on a spit, etc.) pose no problems provided the same area is flushed every time the colony is counted. The use of well-controlled dogs to flush birds within some radius of the center point was also discussed.

   The birds that flush as a result of any method constitutes a working definition of the statistical colony for monitoring purposes, but the proportion of flushing ALTE is largely unknown for all methods. Using cameras or multiple observers, it may be possible for a separate study to estimate the probability of flushing. Ideally, this study could estimate probability of flushing for a range of distances to the flusher, and this could be used to inform the statistical colony definition. For example, assuming 50% of ALTE flush when a human is within M meters, the boundaries of adjacent statistical colonies should be separated by at least $2M$ meters.

2. *Interstitial distance definition*: If gross interstitial distances can be judged during a flush event, it may be possible to define a colony as birds that are within X meters (*X* is likely large) of each other in the air during a flush event. In this case, the idea is to flush as many birds as possible and the flushing method is arbitrary. We recommend that the same flush method be used every visit.

3. *Landmark-based definition*: Assuming suitable vantage points can be permanently established, it is possible to define a statistical colony as flying birds between left and right landmarks. If multiple vantage points are available, all observers must have equal visibility of birds in the sky. Landmarks must be visible during flush counts, generally within an observer's background field of view, and can be features such as rivers, cliffs, beaches, houses, big rocks, roads, etc. It is also possible to establish left and right count boundaries at arbitrary GPS coordinates if the boundary can be visually assessed during flush counting (e.g., temporary orange survey stakes).

B. **How much overcounting is occurring?** It is expected that some unreconciled multiple-observer count model can be formulated to allow for overcounting (the model of Riddle et al. assumes no overcounting) (Riddle et al. 2010). A model that statistically accounts for overcounting is likely necessary because the amount of overcounting is difficult to measure directly and can be significant (but see Section 4.2). If not specifically accounted for in the model, direct flush counts may not be a reliable long-term method unless overcounting can be estimated or separate studies can inform a correction for overcounting (or undercounting). We have difficulty envisioning such a study unless the system is closed (e.g., large aviary) or an intensive method (e.g., drones) has demonstrated high accuracy. If an intensive method is highly accurate and visual counts are taken on the same colonies, it may be possible to develop a calibration curve that removes most of the bias in visual counts.

C. **What are characteristics of the best vantage points, and do they change over time?** Vantage points should be safe for observers and provide unobstructed views of the sky above the colony. Backgrounds should be considered because they affect an observer's ability to detect and count flying birds. For example, when vantage points are at the same elevation as the colony, birds will generally be highlighted against blue sky and clouds. When vantage points are above the elevation of a colony, birds will generally be highlighted against the ground and background vegetation.

D. **What proportion of ALTE flush during the count?** If the proportion of ALTE in the air during any given flush-count is less than 100%, the method must estimate the probability of an individual bird flushing. If the average probability of flushing is π (and we assume flushing is a binomial process), we can likely build this component into the hierarchy of the model of Riddle et al. (2010) to adjust observed counts upward, with a concomitant change in confidence intervals. Adjustments for the flushing percentage should consider nesting stage. The probability of individuals flushing may be affected by nesting status and stage, with individuals tending to flush less during incubation.

Different flushing methods can be used at different colonies if the proportion flushed can be estimated separately. If a common proportion flushed is applied to multiple colonies, the same flushing method should be used. Conversely, if the same flushing method is used at multiple colonies, a common flushing proportion can probably be applied. We recommend that the flush method applied to a statistical colony be fixed for as long as possible, but we also recognize that flush methods will probably change during a long-term study due to advances in understanding and technology (see Letter F below).

At some colonies, it may be possible to wait and count a natural flush event (e.g., after an eagle approach). This method clearly minimizes human disturbances relative to active flush counts. However, natural flushes are less controlled than human flushes and may result in variable proportions of the colony in the air. If researchers are certain that the same proportion of the colony is flushed every time, or the proportion flushed can be accurately estimated each time, counts of natural flushes can be used. Separate studies, perhaps aided by cameras that observe a large fraction of the colony before and after a natural flush event, are likely needed to accurately estimate the natural flushing percentage.

E. **Can we accurately assess the ALTE/ARTE species ratio at mixed colonies during flush counts?** The central question is whether observers can differentiate ALTE and ARTE in flight and obtain either an accurate (i.e., unbiased) count of ALTE or an accurate assessment of the species ratio. Separate and detailed studies are likely necessary to resolve this question. Differentiation between species is likely possible for individual birds, but problematic for large moving groups. Like the proportion of ALTE that flush, it may be possible to build observations on the proportion of ALTE among all terns into the hierarchy of the model of Riddle et al. (2010). If possible, this would adjust mixed-species abundance estimates to reflect ALTE only (and the corresponding confidence intervals). When and if the proportion of ALTE to ARTE changes at a colony, it will be necessary to re-assess the ratio; hence, it will be best to build the assessment into every site visit.

F. **What happens when the flush method changes?** When and if the flushing method changes (e.g., from person to dog, one person to two, straight line to circles), it will be important to accurately estimate the proportion flushing before and after the change. Provided the statistical colony definition does not change and the proportion flushing can be accurately measured, changing flushing method does not create undue problems for estimation.

G. **How much disturbance occurs during flushing?** If flushing causes nest abandonment or provides depredation cues, flush counting will clearly not be workable. The amount of harm done to the colony by presence of the observer and flushing is presumed low but largely unknown at this time and should be assessed. Using cameras or nest checks or band resight information, it may be possible for a separate study to assess survival and other demographic parameters at colonies with and without anthropogenic disturbance.

4.2 Ground-based Photo Counts

The second pilot method attempts to standardize and reduce variation inherent in visual flush counts. Ground-based photo counts are essentially the same as ground-based direct counts (Section 4.1) but observers take multiple photographs in quick succession in an attempt to "freeze" the action and capture flying individuals. After the field visit, individuals in the photos are counted in a mosaic photo that eliminates overlap. Such method will drastically reduce overcount of birds in flight, but may not eliminate the overcount if the same birds cannot be identified in multiple frames (i.e., if overlap is not great enough). This method continues to rely on anthropogenic or natural flushes and hence is dependent on the proportion of flushing individuals and the species ratio at mixed colonies.

4.2.1 Photo-count Protocol

The general protocol for photo-counts is largely the same as the protocol for direct-counts (Section 4.1). Differences are highlighted in **bold**, below.

- Colonies should be visited at least once at a time when nesting activity is anticipated to be high. Nesting activity and colony residency varies from late April to mid-July.
- Visits can be conducted any time during daylight hours.
- Colonies with easy access (e.g., on Kodiak Island's road system) can and should be visited multiple times in a single season. Multiple visits will provide information on nesting phenology and seasonal variation in colony attendance.
- At least **two** people should visit the colony at the same time (one observer and a flusher). Additional observers are recommended.
- If multiple counts are performed in a single day, flushes should be a minimum of 2 hours apart to reduce impacts and to let birds resettle.
- If counting a new colony, observers should pre-scan the colony from a reasonable distance (50 meters) for 10 minutes to determine whether it is a mixed colony.

- At all mixed species colonies, observers should pre-scan the colony from a reasonable distance (50 meters) for 10 minutes and record independent estimates of the proportion of ALTE at the colony.
- Following pre-scan, the **photographer** should position him/herself at **a suitable vantage point** around the colony.
- When ready, the flusher should approach the colony to flush birds.
- **The photographer should attempt to cover the entire statistical colony with a single burst of photos lasting three to five seconds. That is, multiple photos with overlap should be taken so that all in-flight birds appear in the final photo mosaic once.**
- All people should move away from the colony as quickly as possible after the flushing event and otherwise seek to minimize disturbance.

4.2.2 Photo-count Considerations

In addition to the general considerations of Section 3 and the flush-count considerations of Section 4.1, the following photo-count specific considerations apply.

A. **What observation distance is adequate to detect individual birds?** The photo-count method relies on detecting individual birds in stitched photo mosaics of the flush event. Distance from photographer to the colony and resolution of the image will impact our ability to detect and count birds in the photos.

B. **Are ALTE impossible to see in photos with certain backgrounds?** Detection of ALTE in photos will depend upon the photo's background color and complexity. A key question is whether probability of detection drops to an unacceptable level in certain conditions or for certain backgrounds. For example, can ALTE be detected in photos taken on extremely cloudy days?

C. **How much time does it take to process the photos?** After a field visit, photos must be downloaded to appropriate photo-processing software (e.g., Adobe Photoshop or similar), stitched together into a single mosaic to eliminate photo overlap, and all ALTE in the mosaic counted. Some or all of these photo-processing steps can be automated at single-species colonies. At mixed species colonies, it seems unlikely that ALTE and ARTE can be counted separately by humans or an automated process, but this possibility should be investigated.

D. **How reliable are the human or computer-automated counts?** In the long run, an automated ALTE counting algorithm will likely be necessary to efficiently extract counts from a large number of photos. A central question is whether the count, either human or computer-aided or computer-automated, provides an unbiased estimate of the number of ALTE in a photo. Both overcounting and undercounting are possible. To answer this question, it will be important to estimate the probability of detecting and counting an ALTE that is captured in the photo. This can be done by reconciling two or more human counts on a subset of the photo mosaics and allowing other humans and the computer-based algorithms to process the same photos.

4.3. Aerial-based Photo Counts

Pyare et al. (2013) pilot tested use of ArcGIS's built-in Feature Analyst to identify individual terns in photographs taken from fixed-wing aircraft and achieved limited success. Another attempt at some form of aerial photography is warranted for three reasons. First, the method will take photos from a lower altitude (10-30 meters) and should provide enhanced resolution relative to fixed-wing photos. Second, photo processing and detection algorithms have improved significantly since the earlier efforts. Finally, the method does not count flying birds and hence, if successful, would eliminate many questions surrounding flush count methods.

The proposed aerial-photo count method is substantially different than the ground-based direct and photo-count methods. The aerial-photo based count method seeks to count ALTE that are primarily on the ground. Under the aerial-based photo method, regular photos and infrared imagery are obtained from Unmanned Aerial Vehicle (UAV, or “drones”) and ALTE are counted in the imagery either by a technician or photo-recognition algorithm operated in either fully-automated or semi-automated mode.

4.3.1 Aerial-count Protocol

- Colonies should be visited at least once at a time when nesting activity is anticipated to be high. Nesting activity and colony residency varies from late April to mid-July.
- Visits can be conducted any time during daylight hours.
- Colonies with easy access (e.g., on Kodiak Island’s road system) can and should be visited multiple times in a single season. Multiple visits will provide information on nesting phenology and seasonal variation in colony attendance.
- If visiting a new colony, observers should pre-scan the colony from a reasonable distance (50 m) for 10 min to determine whether it is a mixed colony.
- At all mixed species colonies, observers should pre-scan the colony from a reasonable distance (50 m) for 10 min and record independent estimates of the proportion of ALTE at the colony. This estimate of the ALTE/ARTE species ratio will supplement species-specific counts and ratio estimates from the photos.
- The drone pilot should pre-program a set of UAV transects over the colony. For small statistical colonies and those with fixed size, the UAV should generally be programmed to obtain imagery (regular photos or infrared or both) covering 100% of the colony with approximately 10% photo overlap.
- For large statistical colonies (e.g., Black Sand Spit near Yakutat) where it is not possible to photograph the entire statistical colony, the UAV should be programmed to obtain a large spatially balanced sample of photos over the statistical colony (e.g., target 30, 50, or as many photos as possible). No mosaic of the colony is produced

in this case. The analysis will estimate ALTE density at the location of each photo, then interpolate (e.g., kriging or kernel methods) a density map over the entire statistical colony. Some interpolation methods could take into account habitat. Total counts will be inferred from the interpolated maps.

- The UAV should be flown at the lowest possible altitude that does not disturb ALTE on the ground. A pilot drone study carried out near Kenai, Alaska (Magness et al. 2017) indicated that flight altitudes between 10 to 20 meters above ground level do not disturb nesting ALTE. Drone tolerance of non-nesting birds remains to be determined.
- The colony should be disturbed as little as possible.
- After the field visit, imagery should be downloaded from the UAV and, if 100% of the statistical colony was covered, stitched into a single mosaic.

4.3.2 Aerial-count Considerations

In addition to the general considerations of Section 3, the following questions specific to aerial-counts should be considered.

A. **What UAV altitude is best?** The photo-count method relies on detecting individual birds in stitched mosaics. Flight altitude and resolution of the image will impact our ability to detect and count birds in the photos.

B. **What type of UAV is best?** Both rotor-type (copter) and fixed wing drones exist. It is important to consider the most efficient type of drone for surveying each colony and the associated logistics. For example, when rotor-type drones are employed the numbers, cost, and re-charge time of batteries should be considered. When fixed-wing drones are employed, runway length and location must be considered. Due to their ability to hover and collect high-quality orthomosaic photos, we anticipate rotor-type drones will prove most useful for surveying ALTE colonies.

C. **Can and should flying individuals be counted?** We recommend separately counting birds on the ground and in the air so that the proportions of each can be studied. When counting flying birds, higher altitude flights should be employed to minimize birds above the drone, or an estimate of the proportion of flying birds above the drone should be sought, possibly with ground-based photos of the drone and sky above. If an average proportion of birds in the air can be established, drone methods can target birds on the ground and adjusted the count via this proportion.

D. **Can ALTE and ARTE be distinguished in aerial photographs?** Preliminary information (Magness et al. 2017) suggests that ALTE and ARTE can be distinguished from the air, particularly when viewed from an angle (not straight overhead). Hence, it may be necessary to conduct separate UAV flights for counting and assessment of the ALTE/ARTE ratio. The optimal flight altitude for counting may differ from that for species-identification. In particular, infrared imagery may be best suited for counting while regular photography may be best suited for establishing the ALTE/ARTE ratio.

E. **Does the aerial-count method only work in certain habitat?** Detection of ALTE in UAV aerial-photos may depend upon the photo's background color and complexity. A key question is whether probability of detection drops to an unacceptable level in certain habitats (e.g., heavy vegetation). The maximum vegetation height in which ALTE can be reliably detected should be determined.

F. **How much time does it take to process the photos?** After a field visit, photos need to be downloaded to appropriate photo-processing software, stitched together into a single mosaic to eliminate photo overlap, and all ALTE in the mosaic counted. Some or all of the photo processing steps can likely be automated and the reliability of each step in the process should be assessed.

G. **How reliable are the computer automated counts?** Given the large number of photos generated by UAV flights, an automated ALTE counting algorithm will likely be necessary. We recommend a detailed study of the false-positive and false-negative rates, as well as the final overcount and undercount. Such a study can be conducted by comparing human-verified and automated counts on a sample of colony photographs.

H. **Can we obtain permits to fly UAVs in all areas of the state where ALTE colonies exist?** UAVs are regulated by the Federal Aviation Administration (FAA). FAA 14 CFR 107 covers flights for work-related purposes (FAA 2016). In addition, operation of a UAV on US Fish and Wildlife Service National Wildlife Refuges may require a Special Use Permit (SUP) (M. Laker, pers. comm.). In particular, the Alaska Maritime National Wildlife Refuge requires an SUP for UAV operation on refuge lands and an on-site refuge observer may be required. Recipients of Federal grants are responsible for ensuring that all project activities comply with the requirements of the National Environmental Policy Act (NEPA 1969), Section 7 of the Endangered Species Act (ESA 1973), and Section 106 of the National Historic Preservation Act (NHPA 1966). The UAV use must be operated in a manner that is compliant with laws prohibiting disturbance of marine mammals and take of migratory birds. Pilots with Part 107 certification from the FAA should be able to operate UAVs in allowed airspace below 400 feet and within visual line of sight. If colonies are within the airspace of an airport, UAV pilots will be required to coordinate with the airport's air traffic control.

4.4 Call-based Indirect Counts or Indices

Due to recent technological advances in sensitivity, extended duration, and automated detection algorithms, passive acoustic recordings are a viable tool for monitoring seabird species in remote locations (Buxton and Jones 2012). At least one paper (Borker, et al.

2014) reported that call rates are correlated with seabird breeding densities near the acoustic recorders and this type of correlation can provide a reliable metric of seasonal colony phenology. Current call detection methods rely on Deep Neural Networks to classify recorded signals into species-specific calls (McKown et al. 2016). If deployed prior to nest initiation, acoustic recording devices can detect onset and cessation of colony occupancy, disturbance events, and abrupt colony failure (departure).

The indirect abundance estimation method recommended by workshop participants is to deploy acoustic recording devices (e.g., Song Meter units from Wildlife Acoustics, Inc.) at select colonies. Acoustic recordings alone can estimate periods of ALTE-specific occupancy, and may yield an index of colony abundance. Any correlation between the frequency of ALTE-specific calls and either direct counts or nearby nest counts may be used to expand single-visit surveys to either abundance or nest abundance over the full season (Borker et al. 2014).

4.4.1 Call-based Indirect Count Protocol

A general description of the indirect pilot study proposed for summer 2018 includes the following:

- Deploy 15-20 acoustic recording devices prior to onset of nesting (ca. June 1) at selected ALTE colonies across Alaska. Devices will be deployed at colonies representing a range of habitats and expected sizes (or densities), and will include colonies where ALTE direct counts and other monitoring activities are planned. If feasible, some devices will be deployed at remote sites for which no other monitoring protocol is possible and where little information is known.
- At one or more times during the nesting season, personnel will attempt to identify all ALTE nests near the deployed recording devices (ca. within 50 meters) at up to six sites. From these counts, local (to the recorder) nest density can be measured.
- Retrieve the acoustic recorders at the end of the season (ca. July 15).
- Process recordings with acoustic pattern recognition algorithms that exploit the spectro-temporal properties of ALTE calls. These algorithms detect and quantify ALTE-specific acoustic activity rates.
- Relate the estimated number of individuals from direct counts at applicable colonies and over applicable visits to some measure of call frequency or other parameters. This correlation, if found, will allow acoustic recordings alone to estimate abundance throughout the season and at colonies without direct counts.
- Relate nest density (or nest abundance) near the recorders to call frequency or other parameters. At a minimum, such correlation will adjust call frequencies into an index of total colony nest abundance. If a colony has a well-defined area (e.g., an island or spit) and nest density is constant throughout, such correlation will provide an estimate of nest abundance at the colony.

4.4.2 Call-based Indirect Count Considerations

In addition to the general considerations of Section 3, the following questions specific to the proposed indirect method should be considered.

A. **Can ALTE and ARTE calls be distinguished in acoustic recordings?** In theory, it is possible to distinguish ALTE and ARTE calls based on sound-spectral signatures, but this remains to be proven.

B. **How much time does it take to process the recordings?** After device retrieval, recordings need to be downloaded and processed to identify ALTE calls. Most of the sound recording processing steps can likely be automated, but the time required should still be quantified.

C. **How strong is the correlation between call frequency and direct counts?** The strength of correlation between abundance and call frequency (or other parameter) is a primary question. Strong correlations would be necessary to reliably use acoustic recordings to determine estimates of abundance.

D. **How strong is the correlation between call frequency and nest counts?** The strength of correlation between local (to the recorder) nest density and call frequency (or other parameter) is a primary question. Strong correlations and well-defined colony boundaries would be necessary to reliably use acoustic recordings to determine estimates of total colony nest abundance.

4.5 Recommendations for Evaluation of Piloted Methods

Single-colony field methods (the site design) need to be efficient, applicable in a wide range of situations, and provide unbiased estimates of abundance. Most evaluations and comparisons of the proposed methods will happen naturally as researchers apply and refine the methods. Costs and logistics will be much clearer after the pilot work and a comparison of feasibility can happen at that time.

Wherever and whenever possible during summer 2018 work, we recommend conducting all methods on the same colony at the same time, or within a few days of one another. Perhaps it is reasonable at an easy-to-access colony to conduct UAV flights followed immediately by direct- and photo-flush counts. Using data generated by each method, we recommend calculating abundance estimates and confidence intervals to compare the methods. True abundance will not be known in these cases, but we can evaluate whether the methods agree or disagree.

5. Preliminary Recommendations for the Spatial and Temporal Designs

While outcomes of the ALTE Conservation Planning Meeting did not include conclusive recommendations for the spatial and temporal design components of a future ALTE monitoring program, these components were discussed. This section contains bullet lists of preliminary issues and recommendations.

5.1 Spatial Design

Meeting participants did not formalize plans to begin implementing a spatial design. We assume that 2018 surveys will be conducted at sites selected in an opportunistic, non-random fashion. Nonetheless, participants generally identified the following issues that should be considered when designing future spatial sampling plans:

1. Colonies surveyed during previous efforts have come from opportunistic or haphazard sampling of known colony locations or locations surveyed for other objectives (e.g., multi-species surveys). This non-random sampling scheme has restricted the reliability of inference to the broader, unsampled population. Ultimately it will be useful to make inference to all ALTE colonies in the state.
2. To date, there has been little to no directed effort at sampling colonies outside the geographic scope of known breeding areas.

In general, we recommend developing a list of known colonies and fixing this list as the sampling frame. Surveys that visit a probability sample of colonies in the frame will make inference to all colonies in the frame. We recommend that the probability sample of colonies be selected using a spatially balanced sampling method (e.g., BAS or HIP samples) (Robertson et al. 2013, 2017, and 2018). A spatially balanced sample of colonies will ensure that all geographic areas covered by the frame will be represented equally. We also recommend that colonies with long-term unbiased estimates of abundance (legacy sites) be included in the sampling frame, when appropriate.

An important part of the spatial design is the sample frame and how it is updated to include newly discovered colonies. In general, we recommend fixing the frame for a specified number of years (TBD, e.g., 5) and only adding newly discovered colonies at that interval. We envision that new colonies will be discovered through radio and satellite tracking as well as periodic reconnaissance surveys in areas without prior survey effort. During reconnaissance surveys, we recommend including locations not thought to be suitable ALTE habitat. The reconnaissance survey sampling frame could consist of a list of arbitrary spatial units (pixels or coastline segments) or a set of polygons on a map (e.g., Alaska coastline buffered by some amount). In the latter case, points within the polygons would be selected using a probability sampling mechanism (e.g., a spatially balanced algorithm) and colonies within a specified radius of the point would be sought.

5.2 Temporal Design

Regarding seasonal timing, intra-year effort, and rotation of sampling effort among colonies, workshop participants identified the following issues:

1. **When during the season should surveys be conducted?** Annual variability in nesting makes it extremely difficult to pre-determine optimal survey timing. Questions surrounding seasonal timing include: "Should a colony be visited during anticipated peak nesting?" "How early or late in the season can a colony be visited and still produce reliable counts?" Barring additional information from the 2018 pilot season, we lean toward surveys during the incubation stage because counts during that stage will likely be most comparable across colonies. During incubation, colonies appear to have the most consistent use and, on average, a higher proportion of birds on the ground. If surveys rely on flying birds, the latter consideration does not matter. The existing "statewide survey protocol" suggests surveys be conducted between June 10 and July 7 (Pyare et al. 2013). The timing of incubation varies geographically, so survey dates may need to be adapted to local phenology. In this case, auditory recordings from song meters and methods developed during pilot testing may aid in tempering and informing assumptions about the timing of surveys. If auditory recordings are not available or not useful at certain colonies, multiple surveys per year may be necessary to estimate seasonal abundance and to potentially count re-nesting and late-nesting birds.

2. **How many times to survey each year?** The optimum number of times to visit a colony within a year is unclear and will likely be decided after the pilot field season. We generally recommend that colonies be visited only once per year, to spread effort among as many colonies as possible and because transportation to the colonies is expensive. Easy-to-access colonies (e.g., those on the road system) can be visited more often. If auditory recordings are not available, multiple visits will be valuable for establishing phenology and within-season variation.

3. **Should we survey colonies each year?** The optimum rotation of visits among colonies over years is unclear. In general, we recommend a panel design (McDonald 2003), where one panel (i.e., a set of colonies) is revisited each year, and one or more other panels be revisited less frequently.

6. Acknowledgements

We are grateful to the National Fish and Wildlife Foundation Pacific Seabird Program (NFWF-PSP) for providing funding of the 2018 ALTE Conservation Planning Meeting and this report (NFWF-PSP Grant 57885). We acknowledge the Workshop Subcommittee of the ALTE Technical Committee, headed by K. Nesvacil, for organizing the meeting, and we thank J. Caulfield for expert facilitation. We additionally acknowledge M. Goldstein, S. Oehlers, and G. Pendleton for their roles in organizing and facilitating the meeting. D. Lyons, D. Shealer, and L. Welch shared useful insights from their experience researching and managing other tern species.

7. References

Aleutian Tern Conservation Working Group. 2007. *Aleutian Tern collaboration: status, research, and management*. Technical report.

Aleutian Tern Conservation Working Group. 2015. *Summary of information needs and research priorities for Aleutian tern conservation.* Technical report.

BirdLife International (2018) Species factsheet: Aleutian Tern *Onychoprion aleuticus*. Downloaded from http://www.birdlife.org.

Bluso-Demers, J. D., J. T. Ackerman and J. Y. Takekawa. 2010. Colony attendance patterns by mated Forster's terns *Sterna forsteri* using an automated data-logging receiver system. *Ardea*, 98(1):59-65. Available online at: http://www.bioone.org.doi/full/10.5254/078.098.0108.

Borker, A. L., M. W. McKown, J. T. Ackerman, C. A. Eagles-Smith, B. R. Tershy, and D. A. Croll. 2014. Vocal activity as low cost and scalable index of seabird colony size. *Conservation Biology* 28:1100-1108.

Buxton, R. T. and I. L. Jones. 2012. Measuring nocturnal seabird activity and status using acoustic recording devices: applications for island restoration. *Journal of Field Ornithology*, 83:47-60.

Endangered Species Act (ESA). 1973. 16 United States Code (USC) §§ 1531-1544, Public Law (PL) 93-205, December 28, 1973, as amended, PL 100-478 [16 USC 1531 *et seq*.]; 50 Code of Federal Regulations (CFR) 402

Federal Aviation Administration. 2016. Fact Sheet – Small Unmanned Aircraft Regulations (Part 107). Available online at: https://www.faa.gov/news/fact_sheets/news_story.cfm?newsId=20516

Loring, P. H., R.A. Ronconi, L. J. Welch, P.D. Taylor and M. L. Mallory. 2017. Postbreeding dispersal and staging of common and arctic terns throughout the western north Atlantic. *Avian Conservation and Ecology* 12(2):20. Available online at: https://doi.org/10.5751/ACE-01086-120220.

Magness, D. R., T. Eskelin, and M. Laker. 2017. 2017 Census of the Aleutian and Arctic tern colony on Headquarters Lake, Kenai National Wildlife Refuge. U.S. Fish and Wildlife Service unpublished report. 18 pp.

McDonald, T. L. 2003. Review of environmental monitoring methods: survey designs. *Environmental Monitoring and Assessment*, 85(3):277--292, 2003.

McDonald, T. L. 2015. Sampling designs for environmental monitoring. In Bryan F. J. Manly and Jorge A. Navarro Alberto, editors, *Introduction to Ecological Sampling*, chapter 10, pp. 145-165. CRC Press, Boca Raton, Florida.

McKown, M., A. B. Fleishman, and A. D. Earl. 2016. Acoustic surveys for black-capped petrels on Hispaniola and Dominica. Technical report, Conservation Metrics, Inc., Santa Cruz, California.

National Environmental Policy Act (NEPA). 1969. 42 United States Code §4321-4347. January 1, 1970.

National Historic Preservation Act (NHPA). 1966. 16 United States Code §§ 470 *et seq*. October 15, 1966.

North, M. 2013. *Aleutian Tern (Onychoprion aleuticus)*. P. G. Rodewald, ed. The Birds of North America Online. Cornell Lab of Ornithology, Ithaca, New York. Accessed from the Birds of North America Online: https://birdsna-org.bnaproxy.birds.cornell.edu/Species-Account/bna/species/aleter1.

Pyare, S., M. I. Goldstein, D. Duffy, S. Oehlers, N. Catterson, and J. Frederick. 2013. *Aleutian Tern (Onychoprion aleutica) Research in Alaska: survey methodology, migration, and statewide coordination*. Technical report, Juneau, Alaska.

Renner, H. M, M. D. Romano, M. Renner, S. Pyare, M. I. Goldstein, and Y. Artukhin. 2015. Assessing the breeding distribution and population trends of the Aleutian Tern Onychoprion aleuticus. *Marine Ornithology*, 43(2):179-187.

Riddle, J. D., K. H. Pollock, and T. R. Simons. 2010. An Unreconciled double-observer method for estimating detection probability and abundance. *The Auk*, 127(4):841-849.

Robertson, B. L., J. A. Brown, T. L. McDonald, and P. Jaksons. 2013. BAS: balanced acceptance sampling of natural resources. *Biometrics*, 69(3):776--784, 2013.

Robertson, B. L., T. L. McDonald, C. J. Price, and J. A. Brown. 2017. A modification of balanced acceptance sampling. *Statistics and Probability Letters*, 129:107-112.

Robertson, B. L., T. L. McDonald, C. J. Price, and J. A. Brown. 2018. Halton iterative partitioning : spatially balanced sampling via partitioning. *Environmental and Ecological Statistics*, https://doi.org/10.1007/s10651-018-0406-6.

Royle, J. A. 2004. Generalized estimators of avian abundance from count survey data. *Animal Biodiversity and Conservation*, 27(1):375-386.

Taylor, P.D., T. L. Crewe, S. A. Mackenzie, D. Lepage, Y. Aubry, Z. Crysler, G. Finney, C. M. Francis, C. G. Guglielmo, D. J. Hamilton, R. L. Holberton, P. H. Loring, G. W. Mitchell, D. Norris, J. Paquet, R. A. Ronconi, J. Smetzer, P. A. Smith, L. J. Welch, and B. K. Woodworth. 2017. The Motus Wildlife Tracking system: a collaborative research network to enhance the understanding of wildlife movement. *Avian Conservation and Ecology* 12 (1):8. https://doi.org/10.5751/ACE-00953-120108.

Appendix A

Aleutian Tern Conservation Planning Meeting Framework

January 26-27, 2018

This document is a synthesis of ideas from Aleutian tern Technical Committee members, Conservation Planning Meeting Subcommittee members, as well as others. It aims to outline our goal for a 2-day conservation planning meeting to be held January 2018, the primary question we aim to answer, and how we will get there. It also addresses assumptions underlying larger objectives.

Meeting Goal

To facilitate a common understanding among Aleutian tern (ALTE) researchers and managers of alternative sampling and population estimation methods for ALTE, including assumptions, advantages, and limitations, which will lead to consensus recommendations of approaches for future population monitoring.

Primary Meeting Focus Question

How do we best estimate within-year population abundance at the local (i.e., colony) scale at Aleutian tern colonies and Arctic/Aleutian tern colonies?

Estimates of abundance can be divided into two general classes: estimates of population size (i.e., detection and availability probabilities estimated) or estimates of relative abundance. Methods, needed auxiliary data (other than counts), and assumptions differ to produce population estimates vs. relative abundance.

How to determine uniform survey methods to reliably measure abundance at the local scale is our primary focal question due to its fundamental nature in understanding the conservation status of a species or population, and thus its utility from a research and management perspective. Furthermore, as a technical committee, we would like ***future directions*** for Aleutian tern conservation to be based principally on estimates of population size from the colony-level that can be combined to estimate population size at larger spatial and time scales. Estimates of either population size or relative abundance can be used to estimate population trends, either locally or regionally. We intend these estimates of population size and/or trend to be the basis for future conservation planning efforts and management actions.

Considerations for Data Collection

A number of considerations need to be discussed and resolved during the meeting in order to address our focal question, given what is known and unknown about Aleutian terns and their breeding biology.

Discussion will take place during the workshop to determine:

1. the best population variable to measure (e.g., number of: nests, breeding pairs, breeding adults, all adults, juveniles);
2. whether population estimates (need estimates of detection and availability probabilities) or relative abundance (make assumptions about detection and availability) is the optimal measure; and
3. what is the best time scale to assess (e.g., single time points, seasonal totals).

Additional considerations for single-site-year estimates include assessment of the following factors

1. What is a colony? (including “colonies” without nesting?)
2. Habitat as it affects detection probability.
3. Colony configuration and density.
4. Short-term movement (e.g., foraging) as it affects availability to be detected.
5. Intra-annual movement among colonies (for season total estimates).
6. Spatial and temporal variation in nesting phenology (listed here and in Future Directions).
7. Researcher-caused disturbance.

Future Direction - Considerations for Producing Regional and Trend Estimates

1. As stated previously, we would ultimately prefer future directions for Aleutian tern conservation planning to be based on larger spatial and time scales than a single-season at a single colony. As we recognize the need to initially outline the approach to colony abundance assessment, we also recognize the need to be thinking about the future use of these data to ensure we select measures that can be incorporated at larger spatial and time scales. As such we will devote a small section of the workshop to discussing these future directions, including the considerations below. Potentially variable estimation quantities and methods.
2. The effects of unknown colonies on regional estimates.
3. Non-random selection of monitored colonies.
4. Inter-annual variation in colony occupancy and abundance.
5. Appropriate spatial and time scales (for regional abundance and trends, respectively).
6. Spatial and temporal variation in nesting phenology.

Process

A 2-day workshop will be held the 26-27 January, 2018 in order to bring researchers, managers, tern experts, and statisticians together to discuss the primary focal question regarding how to best count Aleutian terns and possible future directions. During this workshop we will achieve the following objectives.

- Share the current state of our knowledge regarding Aleutian tern colony counting methods, with a focus on range of habitats in which surveys have been conducted, and the pros and cons of the methods employed.

- Outside researchers present their work relevant to the meeting question, challenges they have encountered and solutions, or processes they are undertaking to better understand the issues.

- We will hold panel/roundtable discussions to discuss the pros, cons, feasibility, efficiency, and effectiveness of different methods addressing the above considerations for data collection and future directions.

- We will employ the services of a facilitator and statistical consultant to ensure an efficient meeting, and efforts and discussion will be drafted into a monitoring framework with methodology recommendations for the Aleutian tern research and management community.

Summary

Ideally, as a primary conservation question we would like to assess the population abundance and trend for Aleutian tern colonies in Alaska. Realizing the vagile nature of Aleutian tern nesting within and among years, regional differences in nesting phenology, as well as the difficulty of obtaining accurate counts at a colony given the wide range of nesting habitats, at this meeting we will focus on how to best collect colony-level data. We will determine the appropriate parameter to assess and determine the best method(s) to assess that parameter. We will prioritize discussions of how to get appropriate estimates of abundance at individual colonies and, during a short, focused session, consider how to appropriately estimate regional and statewide populations, and determine population trends if feasible.

Workshop Focus Questions and Management Implications (bolded items selected for workshop focus)

What is the population abundance?

Possible parameters assessed: Number of breeding pairs, breeding adults, adults, other age classes

Management Implication: A measure of abundance and/or demography gives us information into robustness of the population to perturbations.

What is their population trend: stable, increasing, or declining?

Possible parameters assessed: Number of nests, pairs, breeding adults, adults

Management Implication: Declining trends could indicate a species of conservation concern, If a declining trend is present, why? What are the limiting factors so we can implement appropriate conservation actions?

What is their reproductive success?

Possible parameters assessed: Breeding propensity, nesting success (hatching), fledging, and recruitment

Management Implication: Identifying bottlenecks at certain stages, and their causes, can aid in development of appropriate conservation actions.

What is their survival?

Possible parameters assessed: Survival for adults, hatch years, other age classes

Management Implication: Identifying bottlenecks at certain stages, and their causes, can aid in development of appropriate conservation actions.

Connectivity among breeding areas

Possible parameters assessed: immigration and emigration among colonies, breeding regions within the state, or between Alaska and Russia

Management Implication: Measures of breeding season connectivity gives us information into robustness of the population to perturbations.

Planning Scale and Colony Composition

Geographic – Entire breeding range, Russian breeding, **Alaskan breeding**, migration, overwintering

Alaskan breeding aggregations – Statewide, Regional, **Colonies**, Dispersed breeding areas

Time – Historic versus **Current** (assess at a later time if we can make inferences to historic counts with current data collections)

Current – **intra-annual variation in colony presence and numbers,** inter-annual variations in colony presence and abundance, determine appropriate time needed for trend estimates

Species – **mixed colonies** versus **Aleutian tern-only colonies** and how to separate counts

Appendix B

Agenda for the Aleutian Tern Conservation Planning Meeting

January 26-27, 2018
Hotel Captain Cook, Resolution Room – Anchorage, Alaska
8:30am-5:00pm each day

Meeting Objective

Facilitate a common understanding among Aleutian tern (ALTE) researchers and managers of alternative sampling and population estimation methods for ALTE, including assumptions, advantages, and limitations, which will lead to recommended methodologies for future population monitoring.

Day 1 - January 26, 2018

8:30 Welcome, Introductions, Meeting Objectives & Agenda

- *Jan Caulfield, Facilitator*
- *Kelly Nesvacil, Alaska Department of Fish & Game (ADFG)*
- *Susan Oehlers, US Forest Service (USFS)*
- *Scott Hall, National Fish and Wildlife Foundation, Pacific Seabird Program*

9:00 Defining the Problem for Aleutian Tern Population Monitoring

Share the current state of our knowledge regarding Aleutian tern colony counting methods, with focus on the range of habitats in which surveys have been conducted, movements of the birds within a season, timing of behaviors, and the pros and cons of the methods employed.

- Previous Statewide Efforts and Synthesis of Current Colony Data – *Heather Renner, US Fish and Wildlife Service (USFWS) (15 min)*
- Kodiak Area Current and Previous Colony Work – *Robin Corcoran, USFWS (15 min)*
- Yakutat Area Previous Colony and Geolocator work – *Mike Goldstein, USFS (15 min)*

9:45 Break

10:00 Continue Defining the Problem

- Yakutat Area Current Colony Work– *Susan Oehlers, USFS (10 min)*
- Southwest Alaska Colony Work– *Kelly Nesvacil, ADFG (10 min)*
  Southwest Alaska and Yakutat PTT data insights– *Susan Oehlers and Kelly Nesvacil (10 min)*

- Summary Presentation – *Grey Pendleton, ADFG (10 min)*

11:30 "Straw Man" Methodology[1]

Introduce "straw man" methodology to use as an initial draft framework for discussion during the work session, with the understanding that it will be modified as a result of participants' expert input, questions, collaborative discussion, and recommendations – *Trent McDonald and Jason Carlisle, Western Ecosystems Technology, Inc. (WEST)*

12:00 Lunch

1:30 Relevant Work with Other Species of Terns

Outside researchers present their work relevant to the meeting question, challenges they have encountered, solutions they have employed, or processes they are undertaking to better understand the issues.

- Arctic and Common Tern – *Linda Welch, USFWS (20 min)*
- Black Tern – *David Shealer, Loras College (20 min)*
- Caspian and Other Terns – *Don Lyons, Oregon State University (20 min)*
- Facilitated discussion – *How this information connects to questions related to ALTE methodology and the "straw man" document (30 min)*

3:00 Break

3:20 New Technologies

Discuss opportunities for use of new technological methods.

- Kenai Drone Surveys – *Dawn Magness, USFWS (15 min)*
- Russian Colony Drone Presentation – *Don Lyons, OSU (10 min)*
- Close Kin Mark-Recapture – *Grey Pendleton, ADFG (15 min)*

4:20 Set the Stage for Day 2 Work Session

On Day 2, WEST will lead a collaborative work session to develop a methodology for ALTE population monitoring, using the "straw man" document as a framework for the discussion. Participants will discuss the pros, cons, feasibility, efficiency, and effectiveness of different methods – and will work with WEST to craft and refine the recommended methodology.

To set the stage for the Day 2 work session, WEST will reprise the "straw man" methodology, discuss how it addresses (or may need to be adjusted to address) points made in Day 1 presentations, and review key questions for discussion on Day 2.

---

[1] From Urban Dictionary – "*Straw man*" refers to an initial document that is expected to be modified by others.

Key Questions to address regarding the Site-Specific Scale:

1. The best population variable to measure (e.g., number of: nests, breeding pairs, breeding adults, all adults, juveniles).
2. Whether population estimates (need estimates of detection and availability probabilities) or relative abundance (make assumptions about detection and availability) is the optimal measure.
3. What is the best time scale to assess within a season (e.g., single time points, seasonal totals).

Considerations related to Site-Specific questions:

- *What is a colony? (including "colonies" without nesting?)*
- *Habitat as it affects detection probability.*
- *Colony configuration and density.*
- *Short-term movement (e.g., foraging) as it affects availability to be detected.*
- *Intra-annual movement among colonies (for season total estimates).*
- *Spatial and temporal variation in nesting phenology.*
- *Researcher-caused disturbance.*

Key Questions to address regarding the Larger Spatial and Temporal Scales:

1. Potentially variable estimation quantities and methods (due to habitat, etc.).
2. The effects of unknown colonies on regional estimates.
3. Non-random selection of monitored colonies.
4. Inter-annual variation in colony occupancy and abundance.
5. Appropriate spatial and time scales (for regional abundance and trends, respectively)
6. Spatial and temporal variation in nesting phenology.

5:00 Adjourn Day 1

Day 2 - January 27, 2018

8:30 Work Session

Collaborative work session using the "straw man" document as a draft framework for developing a recommended methodology for ALTE population monitoring.

12:00 Lunch

1:30 Reconvene Work Session

4:00 Wrap-up – Next Steps, Products, Schedule

Opportunities for outside researchers to share their observations on what we have gained through this process (15 min).

Review results of collaborative work session and next steps / schedule for WEST to produce the draft and final recommended methodology for ALTE population monitoring, with review by workshop participants.

4:50 Closing Comments and Thanks

– ADFG and ALTE Technical Committee

5:00 Adjourn

Appendix C

Aleutian Tern Conservation Planning Meeting Attendance

ATTENDEES	Affiliation	Role
Jan Caulfield	JCC	Facilitator
Trent McDonald	WEST	Consultant/Statistician
Jason Carlisle	WEST	Consultant/Statistician
Linda Welch	USFWS	Invited Expert - Common and Arctic Terns
David Shealer	Loras College	Invited Expert - Black Terns
Don Lyons	OSU	Invited Expert - Caspian Terns
Scott Hall	NFWF	Introductory remarks
Susan Oehlers	USFS	Presenter
Mike Goldstein	USFS	Presenter
Heather Renner	USFWS	Presenter
Robin Corcoran	USFWS	Presenter
Dawn Magness	USFWS	Presenter
Grey Pendleton	ADF&G	Presenter/Statistician
Kelly Nesvacil	ADF&G	Presenter
Bob Gill	USGS	
Aaron Christ	USFWS	
Tory Rhoads	ADF&G	
Jill Tenegeres	OSU/USFWS	
Leah Keeney	USFWS	
Robb Kaler	USFWS	
John Skinner	ADF&G	
Elizabeth Labunksi	USFWS	(Friday only)
Emily Cayer	ADF&G	
Katie Christie	ADF&G	
Kyle James	USFWS	(Friday only)
Chris Krenz	ADF&G	(Friday only)

Appendix D

Aleutian Tern Conservation Planning Meeting Notes

Researchers and managers who work with Aleutian terns (ALTE) in Alaska met January 26-27, 2018 in Anchorage, Alaska, joined by statistical consultants from Western EcoSystems Technology, Inc. (WEST) and researchers who work with different tern species outside of Alaska. The meeting was planned by the ALTE Technical Committee and supported through a grant from the National Fish and Wildlife Foundation. The meeting agenda and attendance list are attached. Presentations made on January 26 regarding Alaska ALTE research, research on other tern species, and new technologies can be viewed at:

http://www.adfg.alaska.gov/index.cfm?adfg=wildlifediversity.aleutian_tern_project.

The following notes summarize discussion during the January 27 work session and list next steps/schedule.

Meeting Objective

Facilitate a common understanding of alternative sampling and population estimates for ALTE, including assumptions, advantages, and limitations, which will lead to recommended methodologies for future population monitoring.

Meeting Outcomes

- WEST will recommend guidelines and protocols for pilot testing three methodologies at selected colonies/sites in 2018, each with the goal of generating an unbiased abundance estimate of ALTE in Alaska.
- WEST's recommendations will address the "considerations" raised in the group's discussion (see notes below).
- WEST will also recommend metrics to use in evaluating the piloted methods – including effectiveness in generating an unbiased abundance estimate for ALTE in Alaska, feasibility, and other relevant factors.
- The group will reconvene in 2019 to evaluate the piloted methods and work toward confirming method(s) that will be used for future population monitoring, including spatial and temporal design of the monitoring program.

Discussion of Monitoring Objectives and Potential Methods

What are we measuring? The objective is to develop and consistently use a method(s) that will generate an unbiased abundance estimate of ALTE in Alaska. This objective led the group to focus on methods that count adult and sub-adult ALTE (both breeding and non-breeding), rather than focusing on counts of nests and/or breeding pairs. The group noted that it is more important (and likely more feasible) to determine population trends, than annual abundance estimates for each colony/site.

Pilot test three methodologies WEST will recommend guidelines and protocols for pilot testing three methodologies in 2018:

1. **Flush counts with photographs of flying adults & sub-adults**

   Method: Flush birds at the colony/site and photograph birds in the air. Count flying birds in the photos. Coincide with double observer counts during flushes and at specified time intervals (direct counts). (See also method #3).

   Considerations

   - What proportion of the birds at the site will flush and be counted vs. proportion that will not flush?
   - What is the best time of season and time of day (relative to solar noon) to use this method? Is it important / possible to standardize the survey window / time of day across sites?
   - Will photo resolution be adequate to count all birds in the air (across distance, etc.)?
   - Will photo resolution be adequate to distinguish ALTE from Arctic terns (ARTE) and other species? Does there need to be a supplemental method to determine the proportion of ALTE/ARTE at the site? (It was cautioned that the species would not likely be intermixed at the same proportion throughout the colony.)
   - Will weather and light conditions confound the birds-in-air photo resolution?
   - What parameters should be used to determine the best vantage point from which to photograph the site?
   - What are we using to flush the birds? People on foot? (Other?) How many needed, per site size and configuration? Any specifications on their spacing, how they flush the birds, etc?
   - Will this method work at large sites, such as Black Sand Spit in Yakutat – or will it need to be adjusted? For example, will large sites need to be partitioned and the method repeated in each sub-area?
   - Note: Some meeting participants were initially not as comfortable with flushing birds from their nests/roosts, but ultimately agreed with testing this methodology. It was also noted in discussion that predators (especially corvids, gulls) may key in on sites after noticing researcher activity and flushed birds.

2. **Photograph ALTE with drones**

   Method: Overfly the colony/site with a drone to photograph birds on the ground and flying below the drone elevation. Suggested that both nesting and non-nesting birds be counted.

Considerations

- What is the best time of season and time of day (relative to solar noon) to use this method? Is it important / possible to standardize the survey window / time of day across sites, or across regions based on regional nesting phenology? (i.e., incubating birds may be the best to survey and coinciding with this timing may provide the most accurate results.)
- Will photo resolution be adequate to distinguish ALTE from ARTE and other species? Does there need to be a supplemental method to determine the proportion of ALTE/ARTE at the site? (Suggested that a lower drone flight could determine species composition, followed by higher flights to count birds. However, it was cautioned that the species would not likely be intermixed at the same proportion throughout the colony.)
- Do we need to account for any birds flying higher than the drone?
- Will photos show birds in denser vegetation?
- Drone cost / availability / logistics
- Pilot cost / training / availability
- Technology dependent – more chance for failed equipment
- Permit issues – e.g., Allowed by NPS? Allowed in Wilderness areas?
- Will this method generate an unbiased abundance estimate?
- Will this method work at large sites, such as Black Sand Spit in Yakutat – or will it need to be adjusted? For example, will large sites need to be partitioned and the method repeated in each sub-area?

3. **Current statewide ALTE survey protocol**

   Method: Consistent implementation of the current statewide "Multiple-Observer Counts of ALTE" guidelines and protocol at each of the pilot sites, to compare with new piloted methods. (As the current survey protocol has been substantially modified by researchers in the past, the best approach would be to conduct double observer counts that coincide with the flush photo counts conducted under method #1.)

How will we evaluate the efficacy and feasibility of the methods? – The group asked WEST to recommend metrics to use in evaluating the piloted methods, so there can be a systematic evaluation of which method(s) are effective in achieving the monitoring objective (unbiased abundance estimate for ALTE in Alaska), are feasible to implement, etc. The group discussed that they need to be able to evaluate whether each method results in a *credible* abundance estimate at the colony/site. Is there any way to check this to increase confidence in the results? (For example, it was suggested that at selected sites the results could be compared with a walk-through count of nests, which researchers feel more confident extrapolating to a count of breeding birds. However, it was also noted that the piloted methods will count all adults, not just breeding birds).

Additional Points of Discussion

- There is interest in having both drone and non-drone options to use. Both will be pilot tested, which will provide information on whether they can be used interchangeably in the future, or if one method is superior to the other.
- There was continued discussion of how to define a "colony" – and this was not resolved at the meeting.
  - The question of how to delineate a "colony" will be important to address in the spatial and temporal design for a long-term monitoring program.
  - It was noted that that the pilot methods will count both breeding and non-breeding birds. Does the term "colony" imply only breeding birds? Would the term "site" be better?
- Regarding when to survey during the day and during the season –
  - How will the pilot methods address the variability of ALTE presence and abundance during the day and during the season?
  - Should the method be pilot tested multiple times during the day and/or season at some sites to get comparative data?
  - Would acoustic recording help?
  - Does existing data provide some guidance re: best seasonal window and/or time of day?
  - Noted that ALTE may abandon the site before incubation – so it may not work to wait to survey during the incubation period.
  - Linda Welch from USFWS Maine Coastal Island Refuges uses nanotags to determine presence at colony and availability.
- Recommendation to use song meters throughout the season to establish dates of presence/absence, nesting phenology, and an index to abundance. (Noted that this is being done for other seabirds, e.g., storm petrels, Forster's tern. Conservation Metrics, Inc. mentioned as one service provider.)
- A long-term methodology will assume just one survey per year at selected sites. However, it would be useful to survey more-accessible sites multiple times in the season to provide additional data (although birds move to other sites in-season, which could result in double-counting).
- There may be promise in using close-kin mark-recapture methodology in the future.[2] The ALTE Technical Committee should consider whether to investigate this method further and provide guidance on whether and how researchers should collect and store DNA samples (e.g., egg membranes, feathers) for future analysis.
- This effort focuses on Alaska ALTE populations. However, it is recognized that an Alaskan abundance estimate does not provide a complete picture for ALTE – which also occur and

[2] Bravington, Mark V., Hans J. Skaug, and Eric C. Andersen, 2016. Close-Kin Mark-Recapture. Statist. Sci. Vol. 31, Number 2(2016), 2590274.

breed in Russia. Alaskan researchers are reaching out to Russian researchers to build collaboration.

Next Steps & Schedule

- Draft recommendations (WEST) – February 28
- Comments submitted to Technical Committee – March 15 (two week review)
    - Technical Committee may ask that each agency consolidate their comments before submitting to WEST.
    - Technical Committee may also consolidate all comments into one submittal to WEST.
- Comments submitted to WEST – March 23
- Final report (WEST) – April 15

The intent is to start pilot testing of these methods in spring/summer 2018. Given the short timeframe, the Technical Committee and researchers may start preparing for pilot testing while the recommendations report is being finalized.