Bridging the gap between disparate phalarope survey methodologies to evaluate population status

Prepared By:

Gabbie Burns, Associate Senior Scientist
Nathan Van Schmidt, Waterbird Program Science Director

San Francisco Bay Bird Observatory
524 Valley Way, Milpitas CA 95035

Prepared For:

Dave Halsing, Executive Project Manager, and Donna Ball, Lead Scientist
South Bay Salt Pond Restoration Project
State Coastal Conservancy

Rachel Tertes, Wildlife Biologist
Don Edwards San Francisco Bay National Wildlife Refuge
U.S. Fish & Wildlife Service

John Krause
California Department of Fish and Wildlife

Amy Larson
California Wildlife Foundation

Laura Cholodenko
California State Coastal Conservancy

FINAL REPORT
April 21, 2023

Burns, G., and N. D. Van Schmidt. 2023. Bridging the gap between disparate phalarope survey methodologies to evaluate population status. San Francisco Bay Bird Observatory. Report prepared for the South Bay Salt Pond Restoration Project.

EXECUTIVE SUMMARY

The South Bay Salt Pond Restoration Project (SBSPRP) is restoring over 15,000 acres of former salt evaporation ponds to a mix of tidal marsh and ponded wetland habitats. These wetlands provide habitat for many waterbirds, including migrating red-necked phalarope *(Phalaropus lobatus*) and Wilson's phalarope (*P. tricolor*). Sustaining baseline population goals for wildlife populations requires understanding how species are responding to restoration actions over time. As part of the Project's Adaptive Management Plan, NEPA/CEQA baseline population levels were calculated based on Salt Pond Survey data from 2005–2007, before major restoration efforts began. If a species group is observed to drop below a sustained threshold or hit a single-year trigger point below that baseline, the Project is committed to evaluating available data and considering targeted management action. In 2017, phalarope counts were 78% below the baseline value, which is well beyond the action trigger of 50% below the baseline of a monthly summer average of 3225 birds.

In response to this observed decline, targeted Phalarope Migration Surveys were piloted in 2019 and have been implemented in 2020–2022 with the goal of better capturing phalarope counts during their brief migration window through the Bay Area and the salt ponds. As these new surveys and the Salt Pond Surveys use different protocols with different levels of detectability, it would be misleading to simply compare raw Phalarope Migration Survey results to the baseline. The goal of this report was to identify data analysis strategies that could best assess the new phalarope migration survey data with respect to the original NEPA/CEQA baseline.

We calculated adjustments to both the Phalarope Migration Survey data and the original baseline to control for differences in survey location, timing, and frequency. In situations where there were multiple options for how to reconcile these differences, we performed each alternative.

The results demonstrate that in every case, the counts from 2020–2022 are more than 50% below the NEPA/CEQA baseline, hitting the management trigger. For future comparisons, we recommend using a revised baseline that includes September surveys to more closely align with the new Phalarope Migration Survey protocol. We also recommend continued Phalarope Migration Surveys until there are sufficient years of survey data to perform an advanced analysis of the rate of population change and differences between the restoration sites and those external to the Project area. At a minimum this requires at least one more year of Phalarope Migration Surveys, to give us a total of three complete years, but the quality of results will improve with further years. Additionally, surveys should be conducted before and after major changes to pond state (e.g., drying or breaching) in order to understand the impact of restoration activities.

INTRODUCTION

In 2003, the U.S. Fish and Wildlife Service (USFWS) and California Department of Fish and Wildlife (CDFW, formerly California Department of Fish and Game) entered into an historic agreement with Cargill Salt to acquire 15,100 acres of salt evaporator ponds in the south San Francisco Bay. The South Bay Salt Pond Restoration Project (SBSPRP) has begun to restore the area to a mix of tidal and ponded habitats while continuing to provide flood risk management and improved public access to many sites.

The SBSPRP has committed to restoring some ponds to tidal marsh, while retaining some pond habitat (as managed ponds) within the project area for waterbirds. Monitoring of waterbird populations is used to assess their effectiveness at achieving these goals. Waterbird surveys were conducted from 2005–2007 to determine population baselines before major restoration activities began. The baselines were used to set thresholds and trigger points for further investigation if the population of a particular species group dropped too low. Certain species groups, such as phalaropes *(Phalaropus spp.)*, were of particular note because their habitat needs are more aligned with the conditions of salt production ponds than restored tidal marsh (South Bay Salt Pond Restoration Project, 2007).

An analysis of salt pond survey data through 2017 (Figure 1) showed a decline of 78% in phalarope counts relative to their NEPA/CEQA pre-restoration baseline (Tarjan 2019a), which hit a management trigger per the Project's adaptive management plan (South Bay Salt Pond Restoration Project 2007). The decline occurred early in the project, almost immediately after the baseline years, and stabilized around 2009. During the years when baseline surveys were conducted from 2005–2007, major restoration activities had not yet begun, except for the tidal restoration of three island ponds (A19–A21). However, initial actions such as reducing salinity in hypersaline ponds to transition them away from salt production were taken as early as summer 2004, and included most ponds by fall 2005 (South Bay Salt Pond Restoration Project 2007). Five ponds (A12–A15, E5–E6) were retained as high salinity ponds. One possible explanation for the observed declines is that these initial actions improved the habitat compared to salt production activities and inflated their abundance during those baseline measurement years. However, the highest count in the SBSPRP area was observed in 2003, before any actions had been taken, and counts were also high outside of the SBSPRP management area in summer 2006 (the first year those sites were surveyed; Fig. 1). This explanation therefore does not comport with available data, and the counts appear to derive from a decline rather than artificially high initial counts.

While the decline in counts likely reflects a decline in phalarope abundance in the current and former salt ponds, the nature of phalarope migration makes it difficult to determine how the magnitude of the observed decline relates to the true magnitude of their population change. Because summer surveys only occurred once per summer, and the timing of counts is highly variable, an alternative explanation could be that previous sampling has to some degree undercounted the true abundance (Tarjan et al. 2019b). To better understand phalarope population trends and pond selection, the SBSPRP Project Management Team requested new monitoring targeted at recording phalaropes during their peak migration (mid-summer through early fall).

This new phalarope migration survey protocol was piloted in 2019, limited surveys were conducted in 2020 due to Covid-19 restrictions, and full surveys were conducted 2021–2022. This protocol has succeeded in generating phalarope counts with higher temporal resolution and lower risk of either double-counting or failing to detect individuals, improving our ability to monitor phalarope population trends (Burns 2021). To inform habitat management decision-making, it is important to understand how the data from these new surveys compares to the baseline established in the SBSPRP adaptive management plan. The methodology differences between the ones used to create the baseline and the new migration surveys mean that data cannot be directly compared between them. The goal of this report was to identify data analysis strategies that could best assess the new phalarope migration survey data with respect to the original NEPA/CEQA baseline.

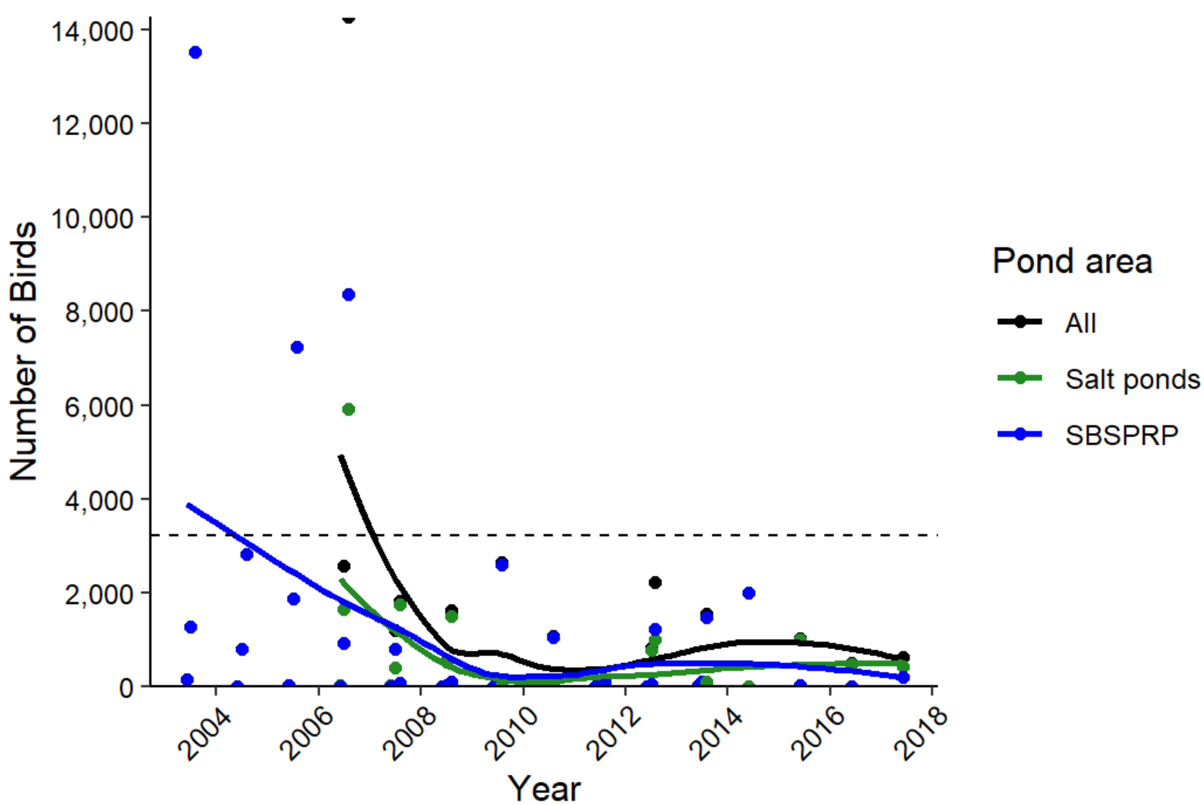

Figure 1. Trends in June, July, and August phalarope abundance from Salt Pond Surveys from the time period when summer surveys were conducted. The dashlined line represents the management trigger threshold.

METHODS

Salt Pond Surveys

The study area included 82 current and former salt ponds in the Santa Clara, Alameda and San Mateo counties of California. The ponds monitored include 25 ponds in the Alviso complex (7404.41 acres), 25 ponds in the Eden Landing complex (4609.04 acres), and 10 ponds in the Ravenswood complex (1487.99 acres), which are part of the SBSPRP. An additional 12 ponds in the Coyote Hills complex (2627.77 acres), 4 ponds in the Dumbarton complex (1287.52 acres), and 6 ponds in the Mowry complex (2940.6 acres) are managed for salt production by Cargill, Inc., but are included in monitoring to understand population trends within the South Bay.

Waterbird censuses of each pond were conducted regularly from 2002 to 2018. At each survey, the total number of individuals of all waterbird species present on each pond was recorded, along with their behavior. Birds in each pond were observed from the nearest drivable levee, using spotting scopes and binoculars. Birds were identified at the species level whenever possible. When species identification was not possible, birds were identified by genus or foraging guild (e.g., phalaropes). Surveys were performed during high tide, defined as a tide of 4.0 feet or greater at the Alameda Creek tide substation (37° 35.7' 122°).

USGS conducted monthly waterbird surveys within the SBSPRP (Eden Landing, Alviso, and Ravenswood complexes) from October 2002 to April 2013, while SFBBO conducted monthly surveys in Cargill-managed ponds (Mowry, Coyote Hills, and Dumbarton) from September 2005 to April 2015. In 2005, summer surveys (June–August) were only conducted at ponds in SBSPRP management area (Alviso, Eden Landing, Ravenswood complexes) and fall surveys (September) were only conducted for salt production ponds (Coyote Hills, Dumbarton, and Mowry complexes). During the years used to calculate the baseline (2005–2007) surveys were conducted once per calendar month by staff from USGS and SFBBO. SFBBO conducted surveys at all 82 ponds during seven 6-week survey periods each year from January 2014 to January 2018, with only one summer survey. Trends for phalaropes were assessed from the summer survey window (Figure 1).

NEPA/CEQA Baseline Calculation

Baselines were calculated using survey data from 2005–2007 from the relevant species group's peak season (i.e., spring, summer, fall, or winter). For phalaropes, this was summer, defined as the months of June, July, and August. For each baseline year, a summer average was calculated by taking the mean of monthly phalarope count totals across all ponds. Counts were included from the SBSPRP ponds (Alviso, Eden Landing, Ravenswood complexes) and the salt production ponds (Coyote Hills, Dumbarton, Mowry complexes). That summer average from each of the three years was then averaged to produce the baseline annual count.

Phalarope Migration Surveys

Phalarope Migration Surveys were conducted at a subset of 24 current and former salt ponds, which identified by SFBBO as comprising >99% of phalarope observations during salt pond surveys between January 2014 and May 2019, plus an additional seven sites outside of the salt pond survey area are surveyed based on high concentrations of phalarope observations recorded in eBird community science data (Figure 2; Tarjan 2019b). The salt ponds monitored for phalaropes by SFBBO include 4 ponds in the Alviso complex, 1 pond in the Coyote Hills complex, 3 ponds in the Dumbarton complex, 10 ponds in the Eden Landing complex, 3 ponds in the Mowry complex, and 3 ponds in the Ravenswood complex. The seven additional sites monitored outside of the salt ponds were Coyote Hills Regional Park, Crittenden Marsh, area Northeast of N1, New Chicago Marsh, Spreckles Marsh, Alviso Marina, and the Sunnyvale Water Pollution Control Plant.

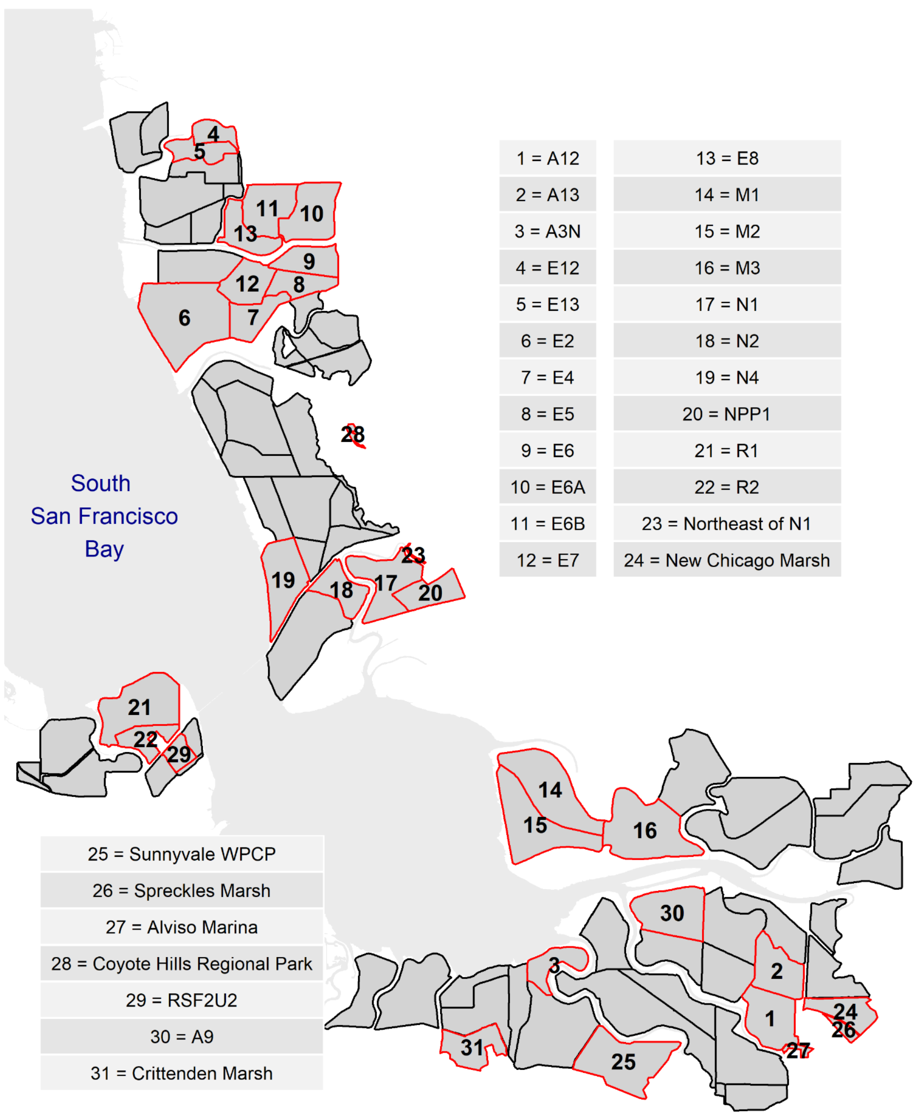

Figure 2. Map of target sites (outlined in red) for Phalarope Migration Surveys.

Surveys are conducted fortnightly from late June through the end of September to capture the peak migration window for each species. The June survey window was added in 2022 to ensure that we captured the leading edge of the Wilson's phalarope (*Phalaropus tricolor)* migration window.

During each survey period, all sites were surveyed as close in time as possible, with most being surveyed simultaneously. Counts were collected by a combination of SFBBO staff and trained community scientist volunteers. Surveyors identified and counted phalaropes at each site from the nearest drivable levee using spotting scopes. Figure 3 shows results from the 2022 season are in Figure 3 (described in Burns et al. 2023).

Statistical modeling to infer missing data

The Covid-19 pandemic shutdown restricted access to sites managed by the US Fish & Wildlife Service in 2020. A small number of surveys were missed in 2021–2022 for other reasons (e.g., active construction). Locations were therefore not sampled in all dates and years. In 2020 only 44% (73/168) of all surveys could be conducted; however in 2021, 99% (167/168) of surveys were conducted, and in 2022, 96% (161/168; Figure 4).

There are several possible approaches to dealing with these missing data. The most basic would be simple mathematical extrapolation: calculate the proportion of surveys that were conducted, X, and multiply the counts by 1/X. In a simplified example: if 6/10 sites were surveyed and a monthly mean of 100 phalaropes were counted, then 0.6 of all sites were surveyed, and the extrapolated monthly mean phalarope count would be $\frac{1}{0.6} * 100 = 167$. However, this approach ignores a crucial aspect of the structure of the data: the missing data are not distributed equally across sites and time (Figure 4). For example, all Alviso and Mowry complex ponds are missing 2020 data. The simple mathematical extrapolation described above would assume that counts from other complexes can be used to infer counts at the unsurveyed sites. We strongly believe that this assumption is not merited.

To account for missing data in a manner that reflects the structure of the data, we fit a statistical model to all years (2002–2022) of phalarope observation data from both Salt Pond Survey and Phalarope Migration Survey datasets. We included Salt Pond Survey data to maximize the available information about phalarope spatial and temporal movements. We included time- and site-specific predictor variables likely to impact the phalarope abundance. We then used the model to infer the values of the missing survey counts. Such inferences are based on all of our information about among-pond and temporal differences in counts, producing abundance estimates that are more likely to be unbiased.

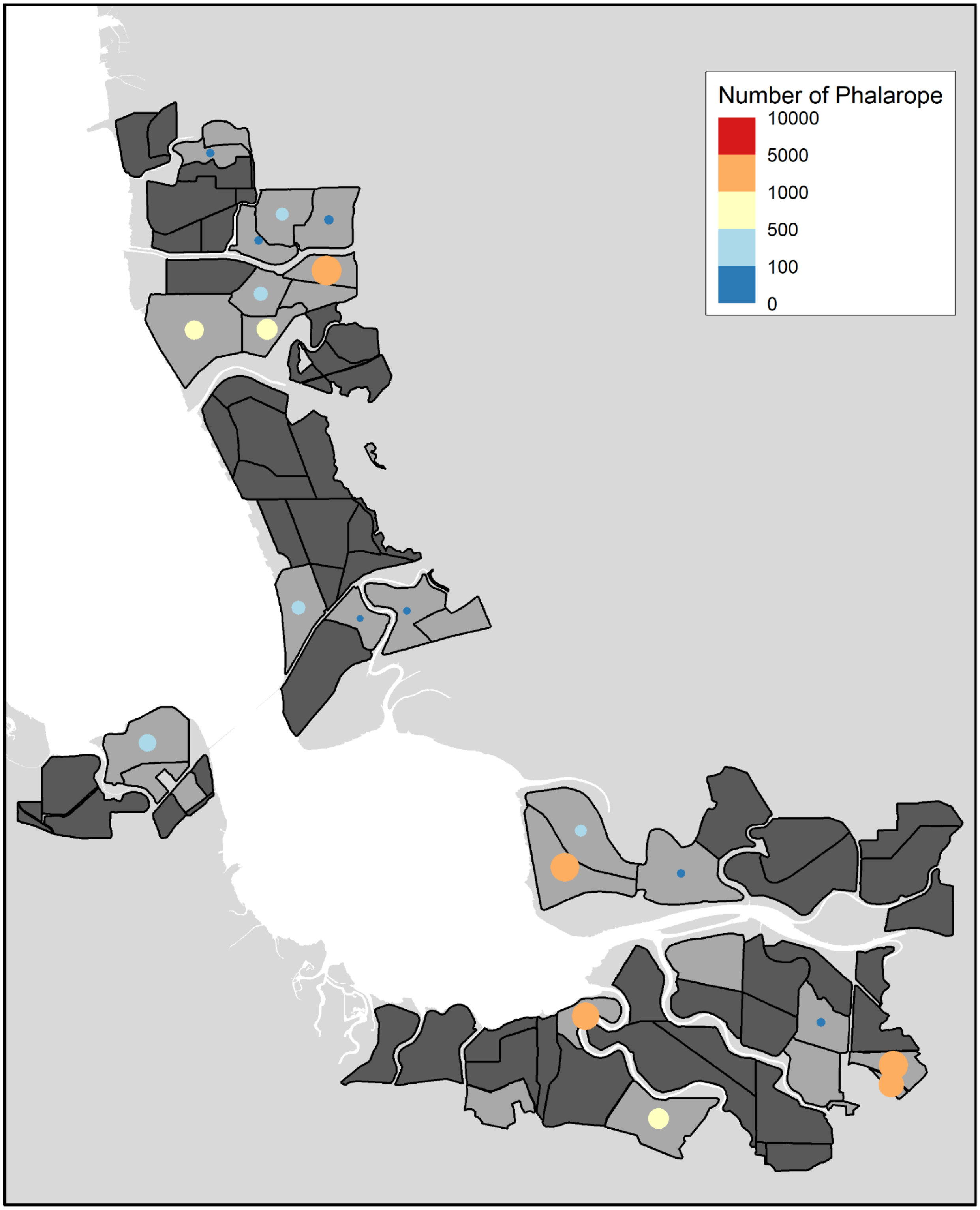

Figure 3. Map with counts of the total number of phalaropes observed during the phalarope migration surveys in 2022. Dots are scaled and colored to represent the number of phalaropes observed; ponds with no dot had zero phalarope observations. Dark gray ponds were not surveyed for Phalarope Migration Surveys in 2022.

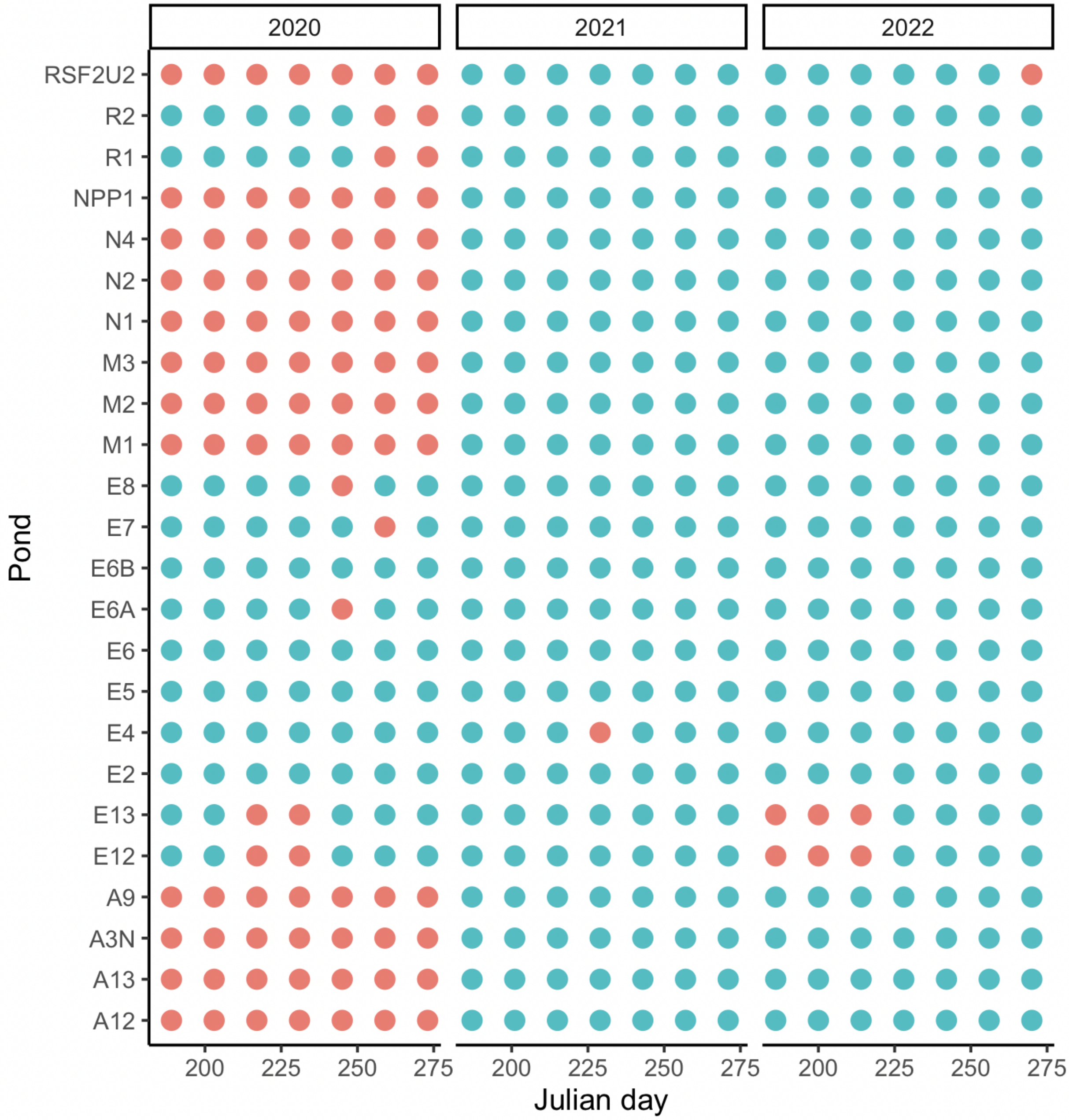

Figure 4. The distribution of missing phalarope migration survey data across years, Julian day, and pond. Turquoise = survey completed, pink = survey not completed (missing data).

We filtered the Salt Pond Survey dataset to include only the months of July, August, and September. These three months cover the overwhelming majority of the fall phalarope migration, which is the main time period covered by the Phalarope Migration Survey dataset, and therefore the time period in which missing data need to be inferred. Of all phalaropes counted in the study area in the past two decades, these three months contain 82% (94,469/115,331). The surrounding months, June and October, contain just 0.03% and 1% of phalaropes counted, respectively. Most of the remaining phalaropes counted (17% of total) were observed in April and May; these represent the spring migration, a separate biological process that is not necessarily informative for predicting fall migration counts.

We used a generalized linear model (GLM), with total phalarope count as the response variable and Julian day, year, and site as fixed effects. We tested a negative binomial error distribution, which is best for datasets with zero inflation (many zeros relative to other values) and overdispersion (high variance relative to the mean), as well as a Poisson error distribution. We used a chi-square test of the model log likelihoods to test which distribution was the best fit to the data, and that model was used for subsequent analyses.

Comparison to the 2005–2007 baseline

The challenge in comparing recent phalarope survey data to the baseline established in 2005–2007 arises from differences in the survey protocols between then and now (Table 1).

One question concerns which months to include in the new baseline. If the goal is to directly compare data between the two, then each dataset must be trimmed to include only months surveyed in *both* the Salt Pond Survey dataset used to calculate the baseline and the Phalarope Migration Survey dataset. This yields a comparison between only surveys during the months of July and August. We term this a *two-month revision*. Alternatively, because the baseline was calculated based on Salt Pond Survey data that was collected year-round, it is possible to redo the

Table 1. Comparison of the survey protocols and the approach to calculating the original baseline value.

Protocol	**Survey frequency**	**Survey synchrony**	**Months (index)**	**Locations**
Salt Pond Survey	1x/month (before 2014), or 1x/6-weeks (2014 on)	Asynchronous	1–12	All current and former salt ponds
Baseline Calculation (2005-2007)	1x/month	Asynchronous	6–8	All[1] current and former salt ponds
Phalarope Migration Survey (2020-2022)	1x/2 weeks	Synchronous	7[2]–9	Subset of salt ponds + additional phalarope sites

[1]In 2005, only the former salt ponds were surveyed during summer months (6-8) and only former salt ponds were surveyed in September.
[2]In 2022 the Phalarope Migration Survey protocols were revised to begin surveying in the second half of June, but we excluded this data because there is only one year of those counts covering only half of the month, rendering them uninformative for the purpose of this comparison.

original calculations to include September survey data from 2005–2007, allowing the three-month period of July–September to be compared to the Phalarope Migration Survey data. This approach has the advantage of using more of the available data and of better measuring the entire fall migration. We term this a *three-month revision*.

There is also a question of which sites to include in calculating the baseline. The new Phalarope Migration Surveys covered only sites believed to support >99% of phalaropes in recent years (Tarjan et al. 2019b), rather than all sites historically monitored to calculate the baseline. Given logistical constraints, this limited subset of sites was necessary to allow for surveys to be completed synchronously . The most direct comparison is therefore to recalculate the baselines *only* for the 24 sites that were surveyed both during the baseline period and during the Phalarope Migration Survey. While conceptually appealing as this compares “apples to apples”, this approach has categorical biases that may be significant drawbacks. If habitat changes since the original baselines were calculated *did* make them unsuitable for phalaropes, these sites would subsequently not be included in the >99% window; excluding data from those sites in the baseline would therefore systematically make the baseline unable to detect the very habitat-driven declines it is supposed to. Conversely, if phalaropes have simply shifted their habitat use to new sites outside of those that were historically monitored, excluding *those* sites from the new monitoring values would make us unable to determine if the population actually has adapted to any changes by switching sites. Therefore, an alternative appealing approach is to simply calculate the baselines on all sites surveyed in each period, even if different sites are surveyed. While this introduces an additional source of error and variability, if a good-faith effort to survey all relevant sites in each period is made, it may still be less biased.

Lastly, a decision is also required with respect to the time period over which the phalarope counts from the Phalarope Migration Surveys should be averaged within each year. Under the Salt Pond Survey protocol, each survey round to visit every pond lasted one month, so the average count per month and per survey round were equivalent. In the Phalarope Migration Survey protocol, a complete survey round occurs every two weeks, which results in two months with two survey rounds each and one month with three survey rounds. One can then either calculate the average total phalarope count per month, or the average total phalarope count per survey round. The former requires first averaging across survey rounds within months, then across months; the latter simply averages across survey rounds irrespective of month. These alternatives are only equivalent if all months have the same number of survey rounds, or if phalarope counts show no pattern of change across months. The former of these conditions is not consistently true, and our past analysis demonstrates clear peaks in the migration timing of Wilson’s and red-necked phalaropes (Tarjan 2019b). We calculated values using both methods to determine whether there was a meaningful difference.

RESULTS

Statistical modeling to infer missing data

A chi-squared test comparing the log likelihoods of the negative binomial model and a Poisson model was highly significant ($p < 0.001$), indicating that the negative binomial model provided a better fit of the data.

A complete set of counts for each Salt Pond site every two weeks from July through September 2020–2022 would contain 504 data points. Our survey dataset contains 401 data points; we used the model to predict the values of the missing 103 data points, most of which (95/103) fell in 2020 due to Covid-19 restrictions. As a general comparison: the predicted count values in 2020 ranged from zero to 974, with a mean of 97.6, while the observed counts for comparable survey round dates and sites from 2021–2022 (the years with nearly complete data) ranged from zero to 4600, with a mean of 72.6. The sum of predicted phalarope counts in 2020 was 9269; the comparable sum from observed counts in 2021 was 9119, and in 2022 was 4536.

Comparison to the 2005–2007 baseline

We calculated both the two-month and three-month alternatives for a revised baseline value (Table 2). For each respective set of sites, the three-month revision was approximately 50% greater than the two-month revision, indicating high counts of phalaropes in September relative to July–August and underscoring the importance of surveying during that month.

We calculated both monthly means and survey round means for the 2020–2022 phalarope migration survey data (Table 3), including the statistically inferred counts. As was the case for the revised baselines, including September counts yielded higher values than did excluding them. The two methods for calculating the mean (monthly vs. survey mean) yielded broadly similar results. In all cases, the 2020–2022 values were considerably (>50%) lower than the comparable revised baseline values (i.e., the baselines calculated using the same months phalarope migration survey value).

Table 2. Results of the several options for calculating a baseline that can be compared to data from the revised Phalarope Migration Survey protocol. All values were calculated from Salt Pond Survey datasets. The three-month original value is reported for comparison with revision values, but was not considered for comparison to Phalarope Migration Survey values.

Calculation approach	Year	Total count	Monthly mean (equal to survey round mean)	Baseline value (mean of monthly means)
Original (Jun–Aug, all historically surveyed sites)	2005	9113*	3038	3225
	2006	16869	5623	
	2007	3039	1013	
Two-month revision (Jul–Aug, only 24 matching sites)	2005	168*	84	2305
	2006	11656	5828	
	2007	2007	1004	
Three-month revision (Jul–Sep, only 24 matching sites)	2005	174*	58	3294
	2006	20411	6804	
	2007	9059	3020	
Two-month-revision (Jul–Aug, all historically surveyed sites)	2005	9110*	3033	3219
	2006	16849	5616	
	2007	3018	1006	
Three-month revision (Jul–Sep, all historically surveyed sites)	2005	9106*	3035	5324
	2006	26810	8937	
	2007	11999	4000	

* 2005 summer surveys (Jun-Aug) were only conducted for ponds in SBSPRP management area (Alviso, Eden Landing, Ravenswood complexes) and 2005 fall surveys (Sep) were only conducted for salt production ponds (Coyote Hills, Dumbarton, and Mowry complexes). Therefore abundances are underestimated compared to years 2006 and 2007.

Table 3. The results of several possible approaches to calculating the average counts in 2020–2022 for comparison to a revised baseline. The baseline value is comparable to both the monthly mean and survey round mean because the baseline protocol involved only one survey round per month, making monthly mean and survey round mean equivalent.

Calculation approach	Year	Mean (% change from baseline)		Comparable revised baseline
		Monthly mean	Survey round mean	
Two-month revision (Jul–Aug, only 24 matching sites)	2020	398 (-82%)	398 (-82%)	2305
	2021	538 (-77%)	597 (-74%)	
	2022	683 (-70%)	801 (-65%)	
Three-month revision (Jul–Sep, only 24 matching sites)	2020	1178 (-64%)	1401 (-57%)	3294
	2021	1462 (-56%)	1373 (-58%)	
	2022	1239 (-62%)	1244 (-62%)	
Two-month revision (Jul-Aug, all historically surveyed sites)	2020	398 (-87%)	398 (-87%)	3219
	2021	538 (-83%)	597 (-81%)	
	2022	683 (-79%)	801 (-75%)	
Three-month revision (Jul-Sep, all historically surveyed sites)	2020	1178 (-77%)	1401 (-74%)	5324
	2021	1462 (-72%)	1373 (-74%)	
	2022	1239 (-77%)	1244 (-77%)	

DISCUSSION

Assessing the Revised Baselines

When the differences in survey frequency is controlled for, it is clear that counts from the Phalarope Migration Surveys are below the NEPA/CEQA baseline at levels that hit the adaptive management action threshold for this species group. This result is robust: it does not change depending on which revised baseline is used (the two-month or the three-month), which set of sites are assessed (all historical sites or matching sites), or which mean calculation is used (monthly mean or survey mean). It is not driven by our inference of missing data, since the years with very little inferred data (2021 and 2022) show the same pattern as the year with more inferred data (2020). The numerical component of the adaptive management action threshold has been decisively met; the attributional component—whether this decline is "likely the result of" restoration actions—is likely to require additional years of Phalarope Migration Survey data to resolve, as well as additional habitat analyses and/or prey items and carrying capacity that might drive observed patterns (Burns et al. 2023).

We recommend recalculating the baseline for the three-month main window of phalarope migration (July through September). It was necessary to recalculate the baseline value because one cannot compare phalarope counts from different times of year without the results being confounded by that difference. The revised baseline values are calculated using the same methods as the original value, so while the value itself is different, the scientific approach of the comparison is unchanged. The revised baseline is also a better measure of phalarope presence in 2005–2007 than the original was, as it includes the phalaropes' true peak migration months.

One set of revised baselines only includes the 24 ponds that are included in the Phalarope Migration Surveys, not all of the current and former salt ponds; this reduced area, if this revised baseline is accepted, would set it apart from the adaptive management baselines set for all other species groups. With respect to detecting declines in phalaropes, the 24-site only baseline is very conservative when it comes to detecting declines because it risks discounting precisely the historical abundance that may have been lost because certain ponds no longer support phalaropes, possibly due to habitat changes resulting from the SBSPRP. A major difference was observed in the original and 24-site only counts in 2005 (Table 2), which was attributable to high abundance at now empty sites (it was not the result of the different temporal window, as only 13 birds were observed in June 2005). Sites that had high abundance in 2005, but are not part of the Phalarope Migration Survey protocol, include:

- A8: 960 phalaropes (total of 3 phalaropes since 2009, most recent detection was in 2015)
- E3C: 1116 (total of 29 since 2009, most recent 2014)
- E5C: 3738 (total of 125 since 2009, most recent 2017)
- E8X: 1804 (total of 4 since 2009, most recent 2010)
- E9: 732 (total of 9 since 2009, most recent 2012)

Notably, E8X and E9 have been restored to tidal action, and A8 was altered to reversibly tidal; it is very plausible the lack of current phalarope use is because of this dramatic habitat shift. This highlights the potential pitfall of discounting historical detections at these sites when calculating a new baseline, because including these sites likely captures an actual local decline due to restoration. While analysis of salt pond data showed that >99% of phalaropes since 2014 have been found in the subset of 24 ponds (Tarjan 2019b), not surveying the other ponds is a matter of expediency given limited funding rather than the ideal scientific approach. Discarding data from no longer occupied sites therefore risks systematically biasing results of monitoring. For these reasons, we believe that all ponds should continue to be counted towards the baseline, even if they are not surveyed in all years because they are no longer utilized as habitat. While a baseline that covers all three months and all surveyed sites is increased compared to the original, the current survey rounds means are also increased, such that the decline remains roughly comparable (Table 3). While not recommended for use as a baseline, we include this very conservative approach to show that a >50% decline that surpasses the adaptive management threshold is still present even after discarding data from no longer utilized sites.

Furthermore, the 2005 contribution to the baseline is artificially low because surveys weren't conducted at any sites outside of the SBSPRP management area, including ponds in the Coyote Hills, Dumbarton, and Mowry complexes that were otherwise counted towards the baseline in 2006 and 2007. If the birds did use all these sites historically, the baseline would be biased lower than reality, and the actual decline would be correspondingly more severe.

Detection probability has also likely increased, a second factor which makes our estimates conservative with respect to estimating declines in the phalarope population. Measured counts of wildlife are the result not only of their abundance but also the detectability of individuals, which ranges 0 (unable to detect any individuals) to 1 (all individuals are detected; Buckland et al. 2008). The Salt Pond Survey protocol is known to be poorly-suited to accurately detecting and counting phalaropes (Tarjan 2019b) due to its low frequency of surveys. Because the Phalarope Migration Survey increases survey frequency by slightly more than 2x, it is likely to detect a higher proportion of all phalaropes than the Salt Pond Survey was. Because the Phalarope Migration Survey likely has a higher detectability value for phalaropes than the Salt Pond Survey did, for any given true population size the Phalarope Migration Survey is likely to yield a higher population size estimate than the Salt Pond Survey (Eq. 1).

Eq.1 *Population size estimate = true population size * detectability* (Buckland et al. 2008)

This means that the comparison of Phalarope Migration Survey counts with the baseline (which was generated from Salt Pond Survey data) is inherently biased in the direction of being *less* likely to detect a decline in phalarope numbers.

That our results still show a major decrease in phalarope counts under all four approaches to revising the baseline despite these biases in the opposite direction underscores the robustness of this result.

Differences that were not controlled for

While recalculating the baseline allowed us to correct many differences between the Salt Pond Survey protocol and the Phalarope Migration Survey protocol, there were several remaining differences that we did not control for. The Salt Pond Surveys occur during high tide, while the Phalarope Migration Surveys occur between 08:00–12:00 regardless of tide level. We believe this is unlikely to have a large impact on our results because none of the surveyed sites are themselves tidal, and anecdotal evidence suggests little use of tidal areas by phalaropes (James Ervin, pers. comm.). If this is true, then our earlier window would most likely favor higher

phalarope counts from the Phalarope Migration Survey by increasing detectability of phalaropes by conducting surveys in the morning when light conditions are favorable and winds are generally reduced.
The protocols also differ in the identity of the surveyors: the Salt Pond Surveys are conducted by SFBBO staff, while the Phalarope Migration Surveys are performed by a mix of staff and volunteers. If these two populations differ in birding ability, that could lead to differences in the accuracy of the phalarope counts. We do not believe this is likely to have a large impact on our results because we only allow volunteers who are skilled birders to participate in the Phalarope Migration Survey, and any new volunteers receive in-the-field training by staff or seasoned volunteers until we are satisfied with their survey ability.

Lastly, the seven sites outside of the historical SBSPRP monitoring area could not be included in the baseline because they were not surveyed, and were therefore also excluded from the new comparison calculations. Because they were not surveyed historically, it is fully unknown whether these represent recent colonizations or historically occupied sites, or whether their abundance has increased or decreased over time. While they cannot be used to determine management triggers, the results from this year that showed an apparent emigration from A13 into these sites highlight that continuing to survey these sites will allow us to better understand the dynamics and movements of phalaropes across the South San Francisco Bay ecosystem.

Recommendations

The Adaptive Management Plan calls for revising baselines as additional data is gathered (South Bay Salt Pond Restoration. Now that we have a more comprehensive understanding of phalarope migration through the Bay Area, we recommend replacing the current baseline with one of the revised baselines presented in this report. Our recommendation is a revised baseline of 5324, which can be compared to a survey round mean of all phalaropes detected within the South Bay. This three-month revised baseline (Jul–Sept) best captures the time span during which phalaropes stopover in the Bay Area, and uses all sites surveyed during the baseline period, rather than filtering the data to include only the set of 24 shared Phalarope Migration Survey sites. As highlighted above, filtering sites risks discounting shifting patterns of relative habitat suitability, and precludes any possibility of recovery of South Bay populations by enhancing or retaining wetlands outside of the SBSPRP footprint as tidal marsh restoration progresses. For comparison with this baseline and calculating the survey round mean rather than the monthly mean, as this avoids any impact of artificial human calendar month distinctions.

In light of the evidence of recent phalarope counts falling well below the management trigger level, we suggest continuing the Phalarope Migration Survey and considering a more complex analysis of the decline once more years of Phalarope Migration Survey data are available. Where the baseline comparison yields a simple binary (*yes the current population is below a management threshold or trigger*, or *no it is not*), collecting more years of data with the improved protocol would open the door to analysis of the rate of population change, as well as comparison of that rate between the current and former salt pond sites and the non-salt pond sites that were introduced as part of the Phalarope Migration Survey. It is quite noteworthy that Wilson's Phalaropes are now most abundant at three sites outside the SBSPRP footprint (New Chicago Marsh, Spreckles Marsh, and Sunnyvale Water Pollution Control Plant). The number of years of data needed for a robust determination could be estimated from a scoping study using data simulation and bootstrapping (Coates et al. 2020). This analysis would permit a better understanding of the cause of any population trend, i.e. whether or not it is specific to the salt ponds and potentially the result of restoration actions. In addition, we recommend continuing to survey in early June to ensure we have a data point before peak migration, because this allows us to be confident that we have accurately captured the true peak if it occurs in late-June/early-July (as it did in 2021).

Given that important sites appear to have significantly changed since the projection's inception and that there is limited funding for fieldwork, we suggest systematic surveys should attempt to keep sampling *effort* consistent from year to year, rather than needing to keep the exact *sites* consistent from year to year. For example, already since the inception of Phalarope Migration Surveys, Crittenden Marsh has been dried for Western Snowy Plover habitat and is unlikely to support phalaropes. Prior to beginning surveys each summer, data from fall/spring Salt Pond Surveys from the previous year should be reviewed to detect red-necked phalaropes in new locations. Furthermore, during Phalarope Migration Surveys, sites should be scanned when passed in transit whenever possible. This could help ensure we can discover and begin monitoring newly utilized sites should the subset of utilized ponds once again shift. Additional funding could further systematize this by devoting one to two days during peak migration to comprehensively scan current and former salt ponds. However, now that monitoring has definitively detected a decline, it may be more fruitful to focus research funding on habitat selection studies rather than minimizing monitoring biases.

ACKNOWLEDGEMENTS

Thank you to the California Wildlife Foundation, the U.S. Fish & Wildlife Service, the California State Coastal Conservancy, and the Santa Clara Valley Water District (Valley Water) for coordinating and funding this work. We would like to thank Rachel Tertes at the Refuge for help with coordinating site access. We would like to acknowledge the South Bay Salt Pond Restoration's Pond Management Working Group that has recommended and implemented many management changes over the years as they relate to specific pond systems. Thank you to Dave Halsing of the South Bay Salt Pond Restoration Project and John Krause with the CDFW Eden Landing Ecological Reserve. We could not have done this work without our dedicated field crews that worked on these surveys over the years.

LITERATURE CITED

Buckland, S., Marsden, S., & Green, R. (2008). Estimating bird abundance: Making methods work. *Bird Conservation International, 18*(S1), S91-S108. doi:10.1017/S0959270908000294

Burns, G. 2021. Phalarope Migration Survey: July – September 2021. Report prepared for the South Bay Salt Pond Restoration Project.

Burns, G., K. LaBarbera, and N. D. Van Schmidt. 2023. Phalarope Migration Survey: June – September 2022. San Francisco Bay Bird Observatory. Report prepared for the South Bay Salt Pond Restoration Project.

Carle, R.D., G. Burns, M. Clapp, K. Caruso, D. House, R. Larson, A. Lewis, A. E. McKellar, J. Neill, M. Prather, J. Reuland, and M. Rubega, and K. Tarr. 2023. Coordinated phalarope surveys at western North American staging sites, 2019-20221. Report of the International Phalarope Working Group.

Coates, P.S., Prochazka, B.G., O'Donnell, M.S., Aldridge, C.L., Edmunds, D.R., Monroe, A.P., Ricca, M.A., Wann, G.T., Hanser, S.E., Wiechman, L.A., and M.P. Chenaille. 2021. Range-wide greater sage-grouse hierarchical monitoring framework—Implications for defining population boundaries, trend estimation, and a targeted annual warning system: U.S. Geological Survey Open-File Report 2020–1154, 243 p., https://doi.org/10.3133/ofr20201154.

South Bay Salt Pond Restoration Project. 2007. Final Environmental Impact Statement/Report. Appendix D: Adaptive Management Plan. http://www.southbayrestoration.org/pdf_files/SBSP_EIR_Final/Appendix%20D%20Final%20AMP.pdf

Tarjan, L.M. 2019a. Trends in waterbird counts in south San Francisco Bay. Report prepared for the South Bay Salt Pond Restoration Project Management Team.

Tarjan, L.M. 2019b. Migration timing and spatial distribution of phalaropes in San Francisco Bay. Report prepared for the South Bay Salt Pond Restoration Project Management Team.