



+ \

Program: BWAS
Grant # 2009-0214

Integrating Avian Datasets for Management, Modeling and Visualization

2010 Annual Report

Submitted by:
Julian Wood, PRBO Conservation Science

Submitted to:
Resources Legacy Fund
555 Capitol Mall, Suite 675
Sacramento, CA 95814

June 29, 2011

Grant Period: October 31, 2009 to October 31, 2011
Reporting Period: October 31, 2009 to December 31, 2010

Introduction

In October of 2009, PRBO Conservation Science initiated the project, “Integrating Avian Datasets for Management, Modeling and Visualization.” The goal of this project is to build a new web portal and database, the Integrated South Bay Avian Database (ISBA-DB) modeled on successful databases previously created at the California Avian Data Center. The South Bay Salt Pond Restoration Project (SBSPRP) needs data on pre- and post-restoration project numbers of birds in the South Bay and in the San Francisco Bay estuary to assess the success of the restoration project for birds.

The original study objectives (Tasks) are as follows:

- 1) Consolidating available bird data within a robust data management system and developing an application framework for data summaries and visualizations useful to managers for tracking and assessing restoration progress.*

Data collected in the past and future will be integrated to provide managers information on restoration progress at multiple scales (pond, complex, South Bay and San Francisco Bay). Data will be integrated and viewed through the California Avian Data Center (CADC; Ballard et al. 2007. www.prbo.org/cadc) an in-use, secure, and well-tested platform that will provide a powerful cost-effective solution to data consolidation and management. Utilities made available through CADC will enable data collected in the past and future to be integrated to provide managers with information on restoration progress at multiple scales (e.g., individual pond,

pond complex, South Bay, and San Francisco Bay). The underlying CADRC relational data model is designed to accommodate different methods and organization of study and project design. It has been successfully used to consolidate point count data and banding data from multiple organizations throughout California (see: <http://data.prbo.org/cadc2/index.php?page=137>). This design will provide easy access to data while ensuring data security, ownership and maintaining data integrity and data sharing policies. We will work with project managers to: 1) populate a customized "ISBA-DB" relational database with San Francisco Bay avian datasets, 2) provide a password-secure repository for San Francisco Bay bird data and data sharing agreements, and 3) provide project managers with a thin client web based data entry, management, graphing/visualization, and analysis tool to facilitate the adaptive management process essential for the success of the SBSPRP.

2) *Integrate carrying capacity modeling with avian population data*

We will integrate pond productivity (chlorophyll) data, macroinvertebrate densities, and waterbird abundance (Takekawa et al. 2005, 2006a) with published information on bird habitat needs to derive a conceptually-based salt pond carrying capacity model for selected species of waterbirds.

3) *Summarize historical waterbird numbers and nesting data*

This effort will concentrate on South Bay bird populations and will use data from the rest of the Bay as comparison. We will review, summarize and synthesize the results of specific studies (both published and unpublished) to describe baseline populations and historical trends for selected species or groups of species. We will estimate baseline nest density, nest success, hatching success, and clutch size for nesting waterbird colonies (avocets, stilts, and Forster's terns) in South San Francisco Bay salt ponds. We will assess habitat characteristics that increase reproductive success. The studies and/or data we will review and summarize include:

- Baseline shorebird and waterfowl numbers in tidal marsh habitat (PRBO)
- Baseline information on waterbird use of salt ponds (PRBO, USGS, SFBBO)
- Bay-wide surveys of shorebirds (PRBO)
- Nesting Snowy Plovers in South Bay salt ponds (SFBBO, FWS).
- Nesting herons and egrets (SFBBO)
- Nesting gulls and terns (SFBBO)
- Clapper Rail, Black Rail abundance and distribution in tidal marsh habitat (PRBO)

4) *Develop a data needs and gaps assessment*

There are many challenges facing managers with the SBSPRP, and it is important to anticipate where baseline data are needed to answer future questions. We will meet with SBSPRP managers and San Francisco Bay Joint Venture (SFBJV) partners to determine the priority data needs for the South Bay. A comprehensive review of information on waterbird use of San Francisco Bay habitats will reveal the following:

- *Gaps in knowledge of bird use of salt ponds and mudflats.* Modeling efforts within this proposal will not only provide information on bird use, but also areas where more research is required.

- *Gaps in available literature:* Our synthesis of available South Bay avian data will provide information on analyses that are possible with currently available data, but have yet to be performed.
- *Gaps in available data.* Our data synthesis and modeling efforts will provide input on which valuable datasets are missing in order to more fully inform SBSPRP about carrying capacity and assess the success of restoration efforts.

Aside from basic population counts, results of the proposed multi-organizational data consolidation effort can be used directly for carrying capacity modeling and “what-if” scenario analyses, including climate change and sea-level models. Additional analytical support will be supported by two proposed graduate fellowships aimed at modeling and predicting waterbird numbers in South Bay ponds and mud flats. ISBA-DB will also inform researchers and wildlife managers on the completeness of the pre-restoration baseline data on waterbird use of the estuary and identify important data gaps needed to evaluate the success of the of the restoration project on a regional scale. Project managers will have the ability to more easily track bird numbers and other parameters using multiple datasets in way that helps the SBSPRP assess restoration progress and achieve their conservation goals.

5) Reporting

PRBO will draft and submit brief quarterly reports (not to exceed 1-2 pages) regarding current activities, project progress towards objectives and milestones, and expected products for the upcoming quarter. PRBO will also submit a more detailed interim report which will summarize results and findings. The final report is described in Task 3 and 4. All reports should be submitted to the SBSP Lead Scientist, RLF, and SCC. PRBO will synthesize and present the findings of this study at relevant meetings, science symposia, and workshops as requested by the SBSP Lead Scientist, RLF, or SCC.

Project Achievements-

1. *Consolidating available bird data within a robust data management system and developing an application framework for data summaries and visualizations useful to managers for tracking and assessing restoration progress.*

a. Interviewed potential ISBA-DB users

We conducted interviews with potential users of ISBA-DB to assess their data needs and find out how they might use the system and what kinds of information they would like to get out of it. Results from these interviews continue to help us tailor data visualizations (e.g., maps, graphs, tables, etc.) and help us design user-friendly data download and upload tools. Potential ISBA-DB users interviewed include: Cheryl Strong, Joy Albertson, and Mike Wolder, USFWS; Laura Valoppi, USGS; and Brian Fulfrost, DC&E. Through these interviews, we learned that most intended users’ top priority is the ability to filter data by species or species group, pond, pond complex or custom aggregations of various ponds, year, month, season and tide. There were multiple requests to view graphs or charts illustrating change in bird use of an area before

and after specific management actions such as levee breaches. The ability to make comparisons among ponds was key to the assessment of various restoration techniques.

b. Subcontracts finalized for San Francisco Bay Bird Observatory and U.S. Geological Survey

SFBBO was subcontracted to prepare and upload avian data for integration into ISBA-DB. Key datasets from SFBBO include Snowy Plover monthly surveys, monthly shorebird surveys and colonial breeding waterbird surveys. Subcontracts were signed with USGS San Francisco Bay Estuary Field Station (USGS-SFBE) to include monthly high and low tide salt pond shorebird surveys and for modeling shorebird carrying capacity. A subcontract was signed with USGS-Davis Field Station to analyze and provide colonial breeding waterbird nesting data for American Avocet, Black-necked Stilt and Forster's Tern from South Bay salt ponds.

c. USGS data proofing

Substantial detailed proofing of USGS-SFBE salt pond monitoring data, including numerous queries to identify inconsistencies in data content and corrections to those data inconsistencies was performed by USGS staff. Before submitting salt pond area search data to ISBA-DB, USGS thoroughly proofed and vetted their data.

d. Developed a data "drop box" website

PRBO developed a web-based data "drop box" for partners to use when submitting data to ISBA-DB. This tool was reviewed by our partners (USGS and SFBBO) and was finalized during the second quarter of 2010. Users were asked to enter information about the data they were submitting (when and where the data were collected, the survey type used, etc.) in a form on the drop box page. To preview the drop box, click <http://www.prbo.org/isbadb/>.

e. Uploaded historical data to ISBA-DB

We uploaded historical avian data from PRBO, USGS and SFBBO. Datasets include PRBO's South Bay Salt Pond Habitat Conversion Modeling Project into ISBA-DB. In addition to bird data, we have identified and integrated data on pond depth, salinity and other site characteristics collected in association with bird data. We integrated six years of shorebird surveys from over 300 sites throughout the Bay into ISBA-DB. These data represent high tide winter shorebird roost counts and can be used to assess shorebird population changes in SBSP project sites relative to Bay-wide changes.

f. Data sharing agreements and data access

Data sharing is a serious concern for the SBSRP research partners. We are addressing these concerns and have been working to develop a data sharing and management structure that reflects the needs of ISBA-DB contributors and users while considering the need to assure data quality and validity, and the privacy rights of land owners, as applicable. PRBO met with SFBBO, USGS and Laura Valoppi, SBSRP Lead Scientist, to discuss data sharing and management issues. The group was invited to comment on a draft data sharing agreement. The data sharing agreement describes how information in ISBA-DB can be shared by a limited and agreed upon list of individuals, henceforth, "core" users (see Appendix A). Although the list of core users has yet to be finalized, the list will include representatives from the SBSRP Management Team (Cheryl Strong, John Krause, Laura Valoppi, and John Bourgeois) and data

contributors from USGS, SFBBO and PRBO. As a result of these meetings and subsequent conversations, it was decided that ISBA-DB users should be presented with the data sharing agreement before accessing data or visualizations.

g. Created application framework and ISBA-DB homepage

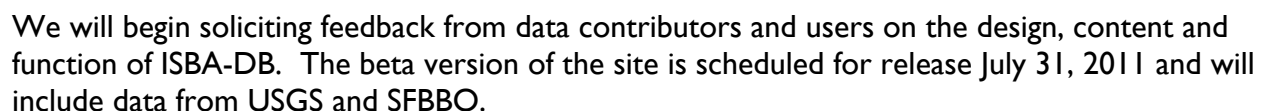
We designed and created the application framework for ISBA-DB which includes the necessary database tables for security/login. The design of the homepage is complete and includes the SBSRP logo, a brief draft summary of the project and buttons that will link to future webpages for maps, data, reports, methods, partners, and data sharing agreements (Figure 1). We are designing the framework for these webpages which will be populated during the first and second quarter 2011.

Figure 1. ISBA-DB homepage screenshot showing SBSRP logo, brief project summary, and links to maps, data, reports (charts and graphs), data, methods, partners, and feedback.



Per the request of data contributors, we ensured that users are made aware of the terms outlined in the data sharing agreement. Users will be required to view the data sharing agreement, scroll down and click “Next” before proceeding to the login page. The user will then be prompted to enter a username and password before viewing any data or visualizations. The framework for this login procedure has been established. Users who do not have a valid login will need to submit a request for approval including a description of who they are and what they plan to use the data for. This request is then subject to approval by project leaders appointed by the data contributors and core users.

Figure 2. ISBA-DB screenshot of map-based avian data summaries application showing shorebird species detected and number of individuals from PRBO data for each species for pond M2 in 1990.



a. Meeting with USGS Vallejo field station

During our meeting with the project partners in June 2010, we discussed with Arianna Brand, USGS-SFBE, the various modeling approaches considered for this task. The methods and final products for this task are still being worked out by USGS.

3. *Summarize historical waterbird numbers and nesting data*

a. Began analysis of nesting data

USGS-Davis analyzed waterbird nesting ecology data from 2005-2010 at sites within the areas of the SBSPRP, including lands managed by the Don Edwards San Francisco Bay National Wildlife Refuge and Eden Landing Ecological Reserve as well as Cargill-owned sites within the Newark salt pond complex. Analyses included nest abundance, nest success, clutch size, nest initiation date, and hatching success. Species analyzed included American Avocet, Black-necked Stilt, and Forster's Tern. Preliminary results indicate that nesting waterbirds are useful indicators for assessing management actions as they can respond quickly to changes in water management and restoration. USGS also began a draft manuscript, "Waterbird nest survival in San Francisco Bay."

4. *Develop a data needs and gaps assessment*

PRBO held a meeting to discuss the approach for addressing the data needs and assessment. PRBO will build and use applications within ISBA-DB to generate summaries of baseline conditions and will identify data gaps during that process. One potential data gap in ISBA-DB that PRBO is working on closing is the lack of 'significant management dates' that relate to pond management events such as increased flows, partial breaches, and to a lesser extent, full breaches. We have begun discussions with SBSPRP managers to determine appropriate dates that relate to these and other important pond management events. Avian conditions before and after these events can be used to elucidate site-specific baseline conditions and avian response to management actions.

In addition, discussions with USGS-SFBE reveal invertebrate data from South Bay mudflats are lacking for conducting in-depth carrying capacity modeling. These and other data needs and gaps will be identified during the carrying capacity modeling and colonial waterbird nesting analyses.

5. *Reporting*

PRBO has submitted 4 quarterly reports covering the period from 31 October 2009 to 30 September 2010. This annual report which summarizes achievements from 31 October 2009 through 29 December 2010 also serves as the fourth quarter 2010 report.

Next Steps

- Sign data sharing agreements with USGS-SFBE and Davis Field Stations.
- Finalize list of core users and establish a process for approving new users

- Populate data summaries and reports applications with USGS and SFBBO data
- Solicit feedback from core users and data contributors on beta release
- Meet with USGS-SFBE to discuss carrying capacity modeling methods and products
- Review avian monitoring reports for Task 4-Data needs and gaps assessment

ISBADB DATA SHARING AGREEMENT

Any data downloaded by core users are subject to any data sharing agreements made between the data contributors and the core users directly. A core user downloading ISBA-DB data should be treated equivalently to a data contributor providing the same data on CD or other media directly to the core user.

- **Restricted--** All data are stored in ISBA-DB primary data warehouses only. Backups of the warehouse are made using persistent data archive techniques. ISBA-DB data managers use all data backup options consistent with the goal of no data loss (backups with periodic data integrity testing). The warehouse serves as the primary archive of all ISBA-DB data, and only applications available to core users will have access to this data. No data is shared with the AKN outside of ISBA-DB.
- **Level 1--**All data are stored in the AKNs primary data warehouse. Backups of the warehouse are made using persistent data archive techniques. AKN data managers use all data backup options consistent with the goal of no data loss (backups with periodic data integrity testing). The warehouse serves as the primary archive of all AKN data, and no applications connect directly to the warehouse. Instead, data from the warehouse are ported to separate data views created specifically to optimize the performance of an application that connects to it. Data owners can specify how their data can be used in the data views, with the option that their data are not exposed to the public at all. For ISBA-DB, only applications available to core users will have access to this data.
- **Level 2--**All of option 1 with the following addition: data can be used in certain publicly available, predefined visualizations (i.e. maps and graphs), but direct access to the data is restricted. For ISBA-DB, only applications available to core users will have access to this data.
- **Level 3--**Data are used in publicly available, predefined visualizations (i.e. maps and graphs). Additionally, the complete BMDE data set is available upon request, subject to approval from the original data provider. The non-core users will be provided access to summary data in ISBA-DB. Summary data may include species lists, aggregated counts at the site (e.g. pond) level, graphs of trends and other visualizations that do not provide raw, record level data or ways of deriving record level information.
- **Level 4--**Data can be used in publicly available, predefined visualizations (i.e. maps and graphs) and also may be available upon request. Additionally, some components of the data are made available to existing bioinformatic efforts (GBIF and ORNIS). These bioinformatic efforts only provide the data "marked-up" to Darwin Core, used to describe primary occurrence (location, date and species for example). The non-core users will be provided access to summary data in ISBA-DB, akin to Level 3 information.
- **Level 5--**Data are used in publicly available, predefined visualizations (i.e. maps and graphs), and are available to existing bioinformatic efforts. Additionally, the complete BMDE data set is available for download directly via the download tool.

To provide the SBSPRP Management Team and other researchers with the most current and comprehensive information, data should be shared whenever possible while considering the tradeoffs between widespread distribution of data sets, the need to assure data quality and validity, and the privacy rights of land owners, as applicable. The general rule is that data should be released as soon as possible, along with documentation that can be used to judge data quality and potential usefulness.

The intellectual investment and time committed to the collection of a data set entitles the investigator to the fundamental benefits of the work. Therefore, publication of descriptive or interpretive results derived immediately and directly from the data is the privilege and responsibility of the investigators who collect the data. This is one of the primary motivations for scientific data collection, and the scientific method is supported by this principle.

Scientists should anticipate that most data collected will eventually become part of the public domain, and recommends that procedures be employed to facilitate this sharing, such as concurrent metadata generation, data verification, data entry, and data backup.

Data for which there are no associated metadata, or which have not been verified, should not be shared. Data that have not been used for peer-reviewed publication should be shared with caution. Data used in publications should be ready for public distribution at the same time as the actual publication. Data relevant to public policy should be shared as quickly and widely as possible. All data that have been shared should be maintained in a fashion that facilitates its continued public accessibility. This includes the need for data versioning, informing users when there have been important changes to public datasets. However, data contributors have no obligation to inform anyone of subsequent changes (such as corrections, additions, deletions) to those data. Data collected on private lands may be subject to more extensive sharing restrictions, depending on agreements between data contributors and collaborators. Some contractual agreements may preclude sharing (or publicly posting on a web or FTP site).

ISBA-DB maintains numerous datasets, including several with thousands of records covering 10 or more years and involving multiple investigators. It is recognized that these datasets can be used for many purposes, either on their own or through combination with other data. Any researcher making substantial use of a dataset should communicate with the investigators who acquired the data, or those to whom responsibility for the data has been entrusted, prior to publication. In most cases these researchers should anticipate that the data collectors or current managers would be co-authors of published results. The ethical and courteous use of data within ISBA-DB is encouraged and will be facilitated by PRBO.

ISBA-DB will maintain a record of all data access and will notify those who access the data that the data may be the intellectual property of the contributor. The appropriate data contributor should be acknowledged in any publication, report, presentation or the like using data that have been collected by the data contributor.

Before any agreements to share data are reached, including posting on a website or FTP server, the specific case must be discussed with the relevant data contributor. The requesting party(s) and the data contributor or authorized person representing the data contributor must sign a

specifically crafted data sharing agreement, memorandum of understanding, or a copy of this Data Sharing Policy acknowledging that all involved agree to abide by its recommendations.

Developed by Grant Ballard, revised per feedback at PRBO's 2003 Science Retreat and with additional revisions and approval by the PRBO Management Team, November 2003; updated per the PRBO Data Committee, October 1, 2007; updated per the Management Team, May 1, 2008; updated per Chief Scientist and Informatics Division review Oct. 8, 2008.

Updated by Julian Wood, June 24, 2010.

PRBO reserves the right to modify, revise or supplement this policy, as it deems appropriate.