Desert Tortoise Data Management
Clark County, Nevada

Final Project Report

Prepared By:
James Cash
Terry Christopher
The Great Basin Institute
7250 S Durango Dr STE 130
PMB #263
Las Vegas, NV. 89113
(702) 433-2600

CONTRACT TITLE: Desert Tortoise Data Management
PROJECT NUMBER: 2021-GBI-2020G
REPORTING DATE: January 31, 2023
PI: Scott Cambrin, Senior Biologist, Desert Conservation Program
DELIVERABLE: Final Project Report (D09)

EXECUTIVE SUMMARY

Project Number: 2021-GBI-2020G
Contract Title: Desert Tortoise Data Management

This report documents data management associated with the Mojave desert tortoise surveys conducted across the range in the desert southwest. In response to the federal listing of the Mojave desert tortoise (*Gopherus agassizii*) as a threatened species, the U.S. Fish and Wildlife Service (FWS) instituted a Mojave Desert Tortoise Range-Wide Monitoring Program to track the population density of Mojave desert tortoises throughout their range. In 2022, the Clark County Desert Conservation Program (DCP) coordinated with FWS and GBI to implement line distance sampling (LDS) to monitor Mojave desert tortoise populations in the eastern Mojave Desert. GBI worked in coordination with FWS and TopoWorks (a private company specializing in GIS and data management) to implement the FWS Data Management Plan (DMP) that governs how the data associated with those surveys are processed and reviewed.

Data records were reviewed from 424 transects across six Tortoise Conservation Areas (TCAs) located in the eastern Mojave Desert. These data included 83 live tortoise and 165 carcass observation records plus 1,092 telemetry observations.

GBI combined the DCP-funded surveys in the six regular monitoring strata with additional desert tortoise surveys funded by the National Park Service (NPS) in nearby (and in some places overlapping) survey strata around Lake Mead National Recreation Area (NRA). GBI also managed data from independent contractor field crews working in southern California on a separate contract with FWS. By concurrently managing data from these separately funded projects GBI and TopoWorks were able to efficiently utilize funds and produce consistent, comparable datasets that allow for integrated analysis and interpretation across the desert tortoise range.

INTRODUCTION

Project Background

The Mojave desert tortoise (*Gopherus agassizii*) is federally listed as threatened under the Endangered Species Act (FWS 1990) and is a priority species for conservation under the Multiple Species Habitat Conservation Plan in Clark County, Nevada (Clark County 2000). The recovery program for desert tortoises requires range-wide, long-term monitoring to determine whether recovery goals are met; specifically, population trends within recovery units need to increase for a period of 25 years to warrant delisting. The purpose of this project is to process and conduct quality assurance/quality control (QAQC) on data associated with line distance sampling (LDS) monitoring efforts across Clark County, Nevada.

Project Description

GBI processed and conducted QAQC on data collected from transect surveys and associated telemetry surveys in Tortoise Conservation Areas (TCAs) located in the eastern Mojave desert in order to insure the highest quality data possible were delivered to DCP and FWS.

The data management process started with testing of the Esri Survey123 electronic data collection forms in partnership with FWS. Field crews were trained in proper use of these forms as well as how to conduct daily data reviews to ensure that data records were complete and matched paper datasheet copies. Data transfers from Survey123 to Microsoft Access were completed regularly during the training and survey periods of the project, followed by systematic data review using manual and automated methods by the GBI data specialist.

Data assessments were provided regularly to field crews in order to address data issues and prevent repeat mistakes. Data submissions were made to DCP at the conclusion of the training phase and weekly during the data collection phase. A final field season database was submitted in July by GBI. GBI and TopoWorks continued with further levels of QAQC and a final QAQC database was submitted by TopoWorks in September.

The methods used to review and process the database followed the FWS Data Management Plan (DMP), which was revised this year by GBI, TopoWorks, and FWS to account for database changes associated with the migration to the Survey123 data collection platform.

Project Goals and Objectives

- Perform QA/QC on data obtained from field crews across the southwest
- Analyze yearly data to provide density estimates for each Critical Habitat Unit

METHODS

Survey123 Form Testing

Data collection forms were created by the FWS on Esri's Survey123 app to facilitate accurate and efficient data collection in the field. GBI thoroughly tested the data forms to ensure the following objectives were met:

- The forms collected the same information requested on the paper datasheets.
- Data entry followed a logical workflow and was user-friendly to ensure efficient, accurate data collection.
- Automated data checks contained within the app worked appropriately to prevent certain types of common data errors while still allowing the full range of possible data entries to be recorded.

The forms were finalized and published for field use following GBI's review.

Data Management Plan, Monitoring Handbook, and Training Material Revisions

GBI worked in partnership with FWS and TopoWorks to revise the DMP, field handbook, and associated training materials to reflect the new workflow associated with the conversion to the Survey123 data collection platform. These updates ensured that field crews and the data management team collected and processed the data appropriately at all stages of the project.

Survey123 to Microsoft Access Transfer

After each data collection exercise during training and weekly during the data collection period the Survey123 records were submitted to ArcGIS Online and paper datasheets were collected and scanned to PDF files. The ArcGIS Online database was downloaded and transferred to a Microsoft Access Collection database using a custom Python script. The data were then transferred from the Collection database to a QAQC database using a custom VBA script that calculated and modified several fields for standardization with previous years' data, as detailed in the DMP. This process allowed the Access database to be reviewed and added to throughout the field season.

Automated Data Review

The data records were reviewed using custom VBA scripts to check for common errors, as detailed in the DMP. These included single-field checks looking for invalid entries (e.g., outside minimum/maximum value ranges) or missing entries, as well as multi-field checks looking for inconsistencies among multiple fields. Data errors were logged in an Error Table to allow easy review and documentation of corrective actions.

In addition to the VBA scripts, the GBI data specialist also developed custom SQL data queries and scripts in program R (R Core Team 2021) in order to either A) more efficiently conduct checks described as manual checks in the DMP, or B) check for new types of errors not addressed in the DMP. Additional R scripts were used to process the data for spatial review in GIS, an additional step that GBI took to ensure data quality that was not required by the DMP.

Manual Data Review
After the automated data checks were completed, the data specialist reviewed the database manually using a combination of sort and filter operations to check for missing, invalid, or contradictory entries. These checks sometimes served as a backup for the automated checks and in other cases looked for types of errors that the automated checks could not address. The GBI data specialist used the manual checks outlined in the DMP as a foundation and built upon that to develop a spreadsheet of specific manual checks associated with each field in the database. This checklist was completed for each weekly data submission.

Additionally, the data specialist uploaded the data in GIS and reviewed walked transect paths in order to detect erroneous GPS locations and give feedback to field teams on their navigation skills and logistical decisions.

Data Corrections
After data errors were identified through manual and automated reviews and logged in the error table, the data specialist attempted to resolve errors by reviewing associated electronic records, paper datasheets, and consulting field teams. If a solution could be identified, the record was corrected and marked as "resolved" in the error table. If the error was not correctable with the information at hand, it was marked as "unresolved" and passed up the QAQC chain for additional review. There were also cases of "exception allowed" error records for cases that met the technical threshold of a data error but that were determined to not be true errors. Finally, some errors records were determined to be erroneously created by malfunctioning automated checks and thus were marked as "script errors."

Data Assessments
GBI provided regular data assessments, modeled after FWS assessments produced in previous years, to field crews so that they were aware of the general progress of the field season as well as data errors or other issues that they should work to avoid in future weeks.

Data Submissions
GBI submitted datasets to DCP as well as finalized QAQC phase 1 pre-season and monitoring season databases. The finalized versions were filtered to exclude non-DCP funded surveys in California and around Lake Mead NRA. TopoWorks submitted the finalized QAQC data products in September.

Data Analysis and Summarization
During the training phase, data from StyroTort training transects were analyzed by GBI and TopoWorks using a combination of existing SQL queries and newly developed R scripts in order to assess each teams' proficiency at maintaining an appropriate detection curve during surveys. The results of these analyses were used to give feedback to teams and resulted in several teams repeating the StyroTort trials as necessary to meet proficiency requirements.

The weekly data assessments included simple statistics covering the number of tortoise observations made and an overall detection distance histogram. These stats gave a preliminary look into the general findings of the project and served as another way to possibly find outliers or anomalies in the dataset.

RESULTS

Objectives

1.) Perform QA/QC on data obtained from field crews across the southwest

- All QAQC phases have been completed.
- Final data products, including spatial and non-spatial data, have been submitted to DCP and FWS.

2.) Analyze yearly data to provide density estimates for each Critical Habitat Unit

- Density estimate analysis has traditionally been carried out by the FWS using standardized processes. GBI does not have access to the scripts used in previous years and so did not conduct a density analysis of the data because doing so without following the exact standardized method used by FWS in previous years may not have produced comparable results and thus would not have been useful for trend analyses. However, GBI did provide preliminary data summaries as part of the data assessments during the 2022 survey season.

QAQC Data Errors

Training Data

- StyroTort: In the StyroTort training dataset there were 181 identified errors, of which 143 were resolved, 1 was an allowed exception, and 36 were unresolved. Common data errors included observations past 25m from the transect line (which were deleted to ease data review/processing, but were not true errors because field crews were instructed to report all observations), incorrect observer name spelling (a problem with the data forms, not the observer's fault), and incorrect tortoise IDs.

- LSTS: The full simulation practice transects at the Large Scale Translocation Site (LSTS) resulted in 308 error records, of which 280 were resolved, 12 were allowed exceptions, and 16 were unresolved. Many of the data errors were the result of the issues with the data collection forms not filling in data fields properly, not unforced errors committed by the observers. The true errors that did exist often had to do with documenting non-standard transects (e.g., listing obstacles).

Transect Data

- Over the course of the 6 week monitoring period there were 382 errors associated with the transect surveys. Of these, 330 were resolved, 47 were determined to be allowed exceptions, and 5 were unresolved. Many errors were the result of technological malfunctions such as incorrect time zones or data fields not filling in correctly under certain circumstances (e.g., child records missing data copied from the parent record if the child record was initiated before the parent record was completely filled out). There were also many errors that were inconsequential such as minor misspellings or inconsistent wording.

G_0 Telemetry Data

- The telemetry technicians logged 135 errors, of which only 1 was left unresolved and 1 was an allowed exception. The other 133 errors were resolved. Common errors included marking iPad GPS grabs as valid even though accuracy was worse than 5m and correcting tort_num for standardization.

Summary Statistics

Field teams from GBI walked 424 transects in the six Tortoise Conservation Areas (TCAs). These data included 10,575 transect waypoints split among 671 transect segments, 71 live tortoises found on transects, 12 live tortoises found opportunistically, 141 tortoise carcasses found on transects, and 24 opportunistic carcass observations. In association with the transect surveys there were 1,092 telemetry observations collected as part of the G_0 correction factor for tortoise visibility. Roughly 51.5% of G_0 telemetry observations were of visible tortoises and 82.5% of telemetry observations were of tortoises in burrows. This is in contrast to a rate of 43.7% of tortoises observed in burrows on the transect surveys.

DISCUSSION

Overall data errors were relatively low given the large volume of data collected, and most errors were either the result of small issues with the data forms or simple typos in data entries. There were few errors that were the result of early substantial misunderstanding of the field protocol or that would have affected data interpretation and analysis if they had gone unchecked. The most common errors that were the result of protocol misunderstanding were associated with documenting non-standard transects, so that should be emphasized in future training.

CONCLUSION

GBI completed all contracted data review and submission deliverables including new types of manual and automated checks to ensure high data quality. TopoWorks finalized and submitted the complete database and related spatial products to DCP and FWS in September 2022.

RECOMMENDATIONS

The GBI data specialist developed new methods of data review in 2022 that went beyond what is laid out in the DMP. Future data specialists should continue to use and build on these tools in order to maintain as efficient and thorough a data review process as possible. Training for field technicians should also emphasize proper documentation of non-standard transects.

LITERATURE CITED

Clark County. 2000. Final Clark County Multiple Species Habitat Conservation Plan and Environmental Impact Statement for Issuance of a Permit to Allow Incidental Take of 79 Species in Clark County, Nevada. September 2000.

FWS (U.S. Fish and Wildlife Service). 1990. Endangered and Threatened Wildlife and Plants; Determination of Threatened Status for the Mojave Population of the Desert Tortoise. Federal Register 55:12178–12191.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.