[*https://doi.org/10.38750/7654321*](https://doi.org/10.38750/7654321)

# DRR Title

### Data Release Report : get this number from Joe DeVivo

#### Jane Doe <https://orcid.org/0000-1111-2222-3333>

NPS Inventory and Monitory Division 1201 Oakridge Dr, Suite 150 Fort Collins, Colorado

#### John Doe

Managed Business Solutions (MBS), a Sealaska Company Contractor to the National Park Service Natural Resource Stewardship and Science Directorate 1201 Oakridge Dr, Suite 150 Fort Collins, Colorado

#### 31 August, 2022

Abstract

Abstract Should go here. Multiple Lines are okay; it’ll format correctly. Pay careful attendtion to non-standard characters, line breaks, carriage returns, and curly-quotes. You may find it useful to write the abstract in NotePad++ or some other text editor and not a word processor (such as Microsoft Word).

The abstract should succinctly describe the study, the assay(s) performed, the resulting data, and their reuse potential, but should not make any claims regarding new scientific findings. No references are allowed in this section.

# Background & Introduction

The Background & Summary should provide an overview of the study design, the assay(s) performed, and the data generated, including any background information needed to put this study in the context of previous work and the literature, and should reference literature as needed. Literature should be cited both in the text and in the References section below using the appropriate [Chicago Manual of Style](https://www.chicagomanualofstyle.org/tools_citationguide/citation-guide-2.html) author-date format.

The section should also briefly outline the broader goals that motivated collection of the data, as well as their potential reuse value. We also encourage authors to include a figure that provides a schematic overview of the study and assay(s) design.

# Data Records (required)

The Data Records section should be used to explain each data record associated with this work (for instance, a data package), including the DOI indicating where this information is stored, and provide an overview of the data files and their formats. Each external data record should be cited.

Tables should be used to support the data records and should clearly indicate the samples and subjects (study inputs), their provenance, and the experimental manipulations performed on each or the protocols for observational data collection (see example Tables below). They should also specify the data output resulting from each data-collection or analytical step, should these form part of the archived record.

### Summary of datasets created (required)

This DRR describes the data package Data Package Title which contains a metadata file and 2 data files. These data were compiled by the National Park Service Biological Resources Division and are available at <https://doi.org/10.57830/12342567> (see Table 1).

**Data Package Title: List of data files**

| File Name | Size | Description |
| --- | --- | --- |
| my\_data.csv | 0.8 MB | This is a short description of my\_data.csv (a good guideline is 10 words or less). |
| my\_data2.csv | 10 GB | This is a short description of my\_data2.csv. |

See Appendix for additional notes and examples.

# Data Quality Evaluation (required)

The Data Quality Evaluation section should present any analyses that are needed to support the technical quality of the dataset. This section may be supported by figures and tables, as needed. *This is a required section*; authors must provide information to justify the reliability of their data. Wherever possible & appropriate, data quality evaluation should be presented in the context of data standards and quality control procedures as prescribed in the project’s quality assurance planning documentation.

**Required elements for this section**

*Stock Text to include:*

The data within the data records listed above have been reviewed by staff in the NPS Inventory and Monitoring Division to ensure accuracy, completeness, and consistency with documented data quality standards, as well as for usability and reproducibility. The *Data Package Title* is suitable for its intended use as of the date of processing (2022-08-31).

*Required Table*

**Description of data quality flags**

| Flag | Definition | Useage |
| --- | --- | --- |
| A | Accepted | columns endining “\_flag” |
| AE | Accepted, estimated | columns endining “\_flag” |
| R | Rejected | columns endining “\_flag” |

Possible content **strongly Suggested to Include**

* Occurrence rates or patterns in data that do not meet established standards or data quality objectives.

Possible content **may include:**

* experiments that support or validate the data-collection procedure(s) (e.g. negative controls, or an analysis of standards to confirm measurement linearity)
* statistical analyses of experimental error and variation
* general discussions of any procedures used to ensure reliable and unbiased data production, such as chain of custody procedures, blinding and randomization, sample tracking systems, etc.
* any other information needed for assessment of technical rigor by reviewers/users

Generally, this **should not include:**

* follow-up experiments aimed at testing or supporting an interpretation of the data
* statistical hypothesis testing (e.g. tests of statistical significance, identifying deferentially expressed genes, trend analysis, etc.)
* exploratory computational analyses like clustering and annotation enrichment (e.g. GO analysis).

# Usage Notes (required)

The Usage Notes should contain brief instructions to assist other researchers with reuse of the data. This may include discussion of software packages (with appropriate citations) that are suitable for analysing the assay data files, suggested downstream processing steps (e.g. normalization, etc.), or tips for integrating or comparing the data records with other datasets. Authors are encouraged to provide code, programs or data-processing workflows if they may help others understand or use the data.

For studies involving privacy or safety controls on public access to the data, this section should describe in detail these controls, including how authors can apply to access the data, what criteria will be used to determine who may access the data, and any limitations on data use.

# Methods

Ideally these methods are identical to the methods listed in the metadata accompanying the data package that the DRR describes.Future versions of this template will pull directly from metadata.

The Methods should cite previous methods under use but also be detailed enough describing data production including a descriptions of the experimental design, data acquisition assays, and any computational processing (e.g. normalization, image feature extraction) that other can understand the methods and processing steps without referring to associated publications. Cite and link to the DataStore reference for the protocol for detailed methods sufficient for reproducing the experiment or observational study. Related methods should be grouped under corresponding subheadings where possible, and methods should be described in enough detail to allow other researchers to interpret the full study.

Specific data inputs and outputs should be explicitly cited in the text and included in the References section below. See the [USGS data citation guidelines](https://www.usgs.gov/data-management/data-citation) for examples of how to cite data in text and in the References section.

Authors are encouraged to consider creating a figure that outlines the experimental workflow(s) used to generate and analyse the data output(s) (Figure 1).



Example general workflow to include in the methods section.

### Data Collection and Sample Processing Methods (optional)

Include a description of field methods and sample processing

### Additional Data Sources (optional)

Provide descriptions (with citations) of other data sources used.

### Data Processing (required if done)

Summarize process and results of any QC processes done that manipulate, change, or qualify data.

### Code Availability (required)

For all studies using custom code in the generation or processing of datasets, a statement must be included indicating whether and how the code can be accessed and any restrictions to access. This section should also include information on the versions of any software used, if relevant, and any specific variables or parameters used to generate, test, or process the current dataset. Actual analytical code should be provided in Appendices.

# References (required)

Provide sufficient information to locate the resource. If the citation has a DOI, include the DOI. If you are citing documents that have unregistered DOIs (such as a data pacakge that you are working on concurrently) still include the DOI. Electronic resources data and data services or web sites should include the date they were accessed.

Bibliographic information for any works cited in the above sections, using the standard *NPS NR Publication Series* referencing style.

ITIS. 2020. Integrated Taxonomic Information System. Available at: <https://www.itis.gov/> (accessed 29 January 2020).

National Park Service (NPS). 2010. Draft Reference Manual RM 66B: Handling Protected Information. National Park Service, Washington, DC. (Available at <https://irma.nps.gov/DataStore/Reference/Profile/2224216>)

National Park Service (NPS). 2019. NPSpecies - the National Park Service biodiversity database. Available at: <https://irma.nps.gov/NPSpecies/> (accessed xx).

Office of Management and Budget (OMB). 2013. Open Data Policy-Managing Information as an Asset. Office of Management and Budget. (Available at <https://digital.gov/open-data-policy-m-13-13/>)

U.S. Fish & Wildlife Service (USFWS). 2019. ECOS Environmental Conservation Online System. Available at: <https://ecos.fws.gov/ecp/> (accessed xx).

U.S. Government Printing Office (GPO). 2013. Making Open and Machine Readable the New Default for Government Information. Executive Order 13642.(Available at <https://www.govinfo.gov/content/pkg/CFR-2014-title3-vol1/pdf/CFR-2014-title3-vol1-eo13642.pdf>)

# Acknowledgements (optional)

The Acknowledgements should contain text acknowledging non-author contributors. Acknowledgements should be brief, and should not include thanks to anonymous referees and editors or effusive comments. Grant or contribution numbers may be acknowledged.

# Appendix A. Code Listing

# The title of your DRR. Should all DRR start with "Data Release Report:"? Should we enforce titles specifically referencing the data package(s) they the report is about?  
title <- "DRR Title"  
  
# Optional and should only be included if publishing to the semi-official DRR series. Contact Joe if you are. If not, leave as NULL  
reportNumber <- ": get this number from Joe DeVivo"  
  
# This should match the Data Store Reference ID for this DRR. Eventually we should be able to pull this directly from the data package metadata.  
DRR\_DSRefID <- 7654321  
  
# Can eventually pull authors from data package metadata, but that assumes data package authors and DRR authors are the same so there will always need to be an option for manually entering author info.  
  
# list the authors.  
authorNames <- c(  
 "Jane Doe",  
 "John Doe"  
)  
  
# List author affiliations. The order of author affiliations must match the order of the authors in AuthorNames. Note that the entirety of each affiliation is enclosed in quotations. Line breaks are indicated with the tag <br>. Do not worry about indentation or word wrapping.  
authorAffiliations <- c(  
 "NPS Inventory and Monitory Division<br>  
1201 Oakridge Dr, Suite 150<br>  
Fort Collins, Colorado",  
 "Managed Business Solutions (MBS), a Sealaska Company<br>  
Contractor to the National Park Service<br>  
Natural Resource Stewardship and Science Directorate<br>  
1201 Oakridge Dr, Suite 150<br>  
Fort Collins, Colorado"  
)  
  
# List the ORCID iDs for each author in the format "xxxx-xxxx-xxxx-xxxx". If an author does not have an ORCID iD, specify NA (no quotes). For more information on ORCID iDs and to register an ORCID iD, see https://www.orcid.org. The order of the ORCID iDs must match the order of authors in AuthorNames.In this example, Jane Doe has an ORCID iD but John Doe does not.  
authorORCID <- c(  
 "0000-1111-2222-3333",  
 NA  
)  
  
DRRabstract <- "Abstract Should go here. Multiple Lines are okay; it'll format correctly. Pay careful attendtion to non-standard characters, line breaks, carriage returns, and curly-quotes. You may find it useful to write the abstract in NotePad++ or some other text editor and not a word processor (such as Microsoft Word).\n\n  
  
The abstract should succinctly describe the study, the assay(s) performed, the resulting data, and their reuse potential, but should not make any claims regarding new scientific findings. No references are allowed in this section."  
  
# DataStore reference ID for the data package associated with this report. You must have at least one data package.Eventually, we will automate importing much of this information from metadata.  
dataPackageRefID <- c(12342567)  
  
# Must match title in DataStore and metadata  
dataPackageTitle <- "Data Package Title"  
  
# Must match descriptions in the data package metadata  
dataPackageDescription <- "Short title for data package1"  
  
# generates your data package DOI based on the data package DataStore reference ID. This is different from the DRR DOI! No need to edit this.  
dataPackageDOI <- paste0("https://doi.org/10.57830/", dataPackageRefID)  
  
# list the file names in your data package. Do NOT include metadata files.  
dataPackage\_fileNames <- c(  
 "my\_data.csv",  
 "my\_data2.csv"  
)  
  
# list the approximate size of each data file. Make sure the order corresponds to the order of of the file names in dataPackage\_fileNames  
dataPackage\_fileSizes <- c("0.8 MB", "10 GB")  
  
# list a short, one-line description of each data file. Descriptions must be in the same order as the filenames.  
dataPackage\_fileDescript <- c(  
 "This is a short description of my\_data.csv (a good guideline is 10 words or less).",  
 "This is a short description of my\_data2.csv."  
)  
  
# Write parameters to file for consistent reuse across report and data packages  
save(title, reportNumber, DRR\_DSRefID, authorNames, authorAffiliations, authorORCID, DRRabstract, dataPackageRefID, dataPackageTitle, dataPackageDescription, dataPackageDOI, dataPackage\_fileNames, dataPackage\_fileSizes, dataPackage\_fileDescript, file = here::here("output", "temp", "reportParameters.Rdata"))  
RRpackages <- c("markdown", "rmarkdown", "pander", "knitr", "fontawesome", "yaml")  
  
inst <- RRpackages %in% installed.packages()  
if (length(RRpackages[!inst]) > 0) {  
 install.packages(RRpackages[!inst], dep = TRUE, repos = "https://cloud.r-project.org")  
}  
lapply(RRpackages, library, character.only = TRUE)  
  
# \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
# Now repeat for packages used in the analyses  
pkgList <- c("devtools", "RODBC", "remotes", "dplyr", "ggplot2", "plotly")  
  
inst <- pkgList %in% installed.packages()  
if (length(pkgList[!inst]) > 0) {  
 install.packages(pkgList[!inst],  
 dep = TRUE,  
 repos = "https://cloud.r-project.org"  
 )  
}  
  
lapply(pkgList, library, character.only = TRUE, quietly = TRUE)  
date <- format(Sys.time(), "%d %B, %Y")  
cat("#", title, "\n")  
if (!is.null(reportNumber)) {  
 subtitle <- paste0("Data Release Report ", reportNumber)  
 cat("###", subtitle)  
}  
contact <- data.frame(authorNames, authorAffiliations, authorORCID)  
  
for (i in seq\_along(authorNames)) {  
 curr <- contact[i, ]  
 cat("####", curr$authorNames, " ")  
 if (is.na(curr$authorORCID)) {  
 cat("\n")  
 }  
 if (!is.na(curr$authorORCID)) {  
 # Consider re-write to use the actual downloaded icon and not fontawesome.For some reason fill isn't rendering in chrome (but does in edge). Maybe move fill to CSS?  
 orc <- paste0("https://orcid.org/", curr$authorORCID)  
 cat({{ orc }})  
 }  
 cat("\n")  
 cat(curr$authorAffiliations, "\n\n")  
}  
  
filelist <- data.frame(dataPackage\_fileNames, dataPackage\_fileSizes, dataPackage\_fileDescript)  
  
knitr::kable(filelist, caption = paste0("\*\*", dataPackageTitle, ": List of data files\*\*"), col.names = c("File Name", "Size", "Description"), format = "pandoc")  
flags<-c("A", "AE", "R")  
def<-c("Accepted", "Accepted, estimated", "Rejected")  
app<-rep("columns endining \"\_flag\"", 3)  
data\_flags<-data.frame(flags, def, app)  
  
knitr::kable(data\_flags, caption = "\*\*Description of data quality flags\*\*", col.names=c("Flag", "Definition", "Useage"), format="pandoc")  
include\_graphics("output/temp/example\_DRR/figures/ProcessingWorkflow.png")  
  
sessionInfo()  
Sys.time()

# Appendix B. Session and Version Information

## R version 4.2.0 (2022-04-22 ucrt)  
## Platform: x86\_64-w64-mingw32/x64 (64-bit)  
## Running under: Windows 10 x64 (build 19042)  
##   
## Matrix products: default  
##   
## locale:  
## [1] LC\_COLLATE=English\_United States.utf8   
## [2] LC\_CTYPE=English\_United States.utf8   
## [3] LC\_MONETARY=English\_United States.utf8  
## [4] LC\_NUMERIC=C   
## [5] LC\_TIME=English\_United States.utf8   
##   
## attached base packages:  
## [1] stats graphics grDevices utils datasets methods base   
##   
## other attached packages:  
## [1] plotly\_4.10.0 ggplot2\_3.3.6 dplyr\_1.0.9 remotes\_2.4.2   
## [5] RODBC\_1.3-19 devtools\_2.4.4 usethis\_2.1.6 yaml\_2.3.5   
## [9] fontawesome\_0.3.0 knitr\_1.39 pander\_0.6.5 rmarkdown\_2.15   
## [13] markdown\_1.1   
##   
## loaded via a namespace (and not attached):  
## [1] Rcpp\_1.0.9 here\_1.0.1 tidyr\_1.2.0 png\_0.1-7   
## [5] prettyunits\_1.1.1 ps\_1.7.1 rprojroot\_2.0.3 digest\_0.6.29   
## [9] utf8\_1.2.2 mime\_0.12 R6\_2.5.1 evaluate\_0.16   
## [13] highr\_0.9 httr\_1.4.3 pillar\_1.8.1 rlang\_1.0.4   
## [17] lazyeval\_0.2.2 rstudioapi\_0.13 data.table\_1.14.2 miniUI\_0.1.1.1   
## [21] callr\_3.7.1 urlchecker\_1.0.1 stringr\_1.4.1 htmlwidgets\_1.5.4  
## [25] munsell\_0.5.0 shiny\_1.7.2 compiler\_4.2.0 httpuv\_1.6.5   
## [29] xfun\_0.31 pkgconfig\_2.0.3 pkgbuild\_1.3.1 htmltools\_0.5.3   
## [33] tidyselect\_1.1.2 tibble\_3.1.7 viridisLite\_0.4.0 fansi\_1.0.3   
## [37] crayon\_1.5.1 withr\_2.5.0 later\_1.3.0 grid\_4.2.0   
## [41] jsonlite\_1.8.0 xtable\_1.8-4 gtable\_0.3.0 lifecycle\_1.0.1   
## [45] magrittr\_2.0.3 scales\_1.2.0 cli\_3.3.0 stringi\_1.7.8   
## [49] cachem\_1.0.6 fs\_1.5.2 promises\_1.2.0.1 ellipsis\_0.3.2   
## [53] generics\_0.1.3 vctrs\_0.4.1 tools\_4.2.0 glue\_1.6.2   
## [57] purrr\_0.3.4 processx\_3.7.0 pkgload\_1.3.0 fastmap\_1.1.0   
## [61] colorspace\_2.0-3 sessioninfo\_1.2.2 memoise\_2.0.1 profvis\_0.3.7

## [1] "2022-08-31 08:49:06 MDT"