Coastal Wetlands Indicator Quantile Report

SEACAR Analysis

Last compiled on 24 October, 2023

Contents

Overview

Purpose

The purpose of the indicator quantiles is to flag records that are "unusual" relative to all of the data in the DDI for a given indicator in order to facilitate QA/QC. They are not used to filter any of the data for SEACAR analyses, and the presence of a LowerQuantile or UpperQuantile flag on a DDI record alone does not necessarily indicate there is any issue with the record (neither does the absence of a LowerQuantile or UpperQuantile flag necessarily mean that a data record is correct).

Relevant file locations

Current values can be found in the "LowQuantile" and "HighQuantile" columns of the "Ref_Parameters" worksheet.

The R script described below and the output file can be found in the *FloridaSEACAR IndicatorQuantiles* repository on GitHub:

- https://github.com/FloridaSEACAR/IndicatorQuantiles

Process steps

IQ_Report_Render.R & IQ_Report.Rmd

1. The *IQ_Report_Render.R* script lists all files in a given directory and filters it to a list of DDI exports to evaluate considering a list of parameters to skip (user-defined).

2. User sets the desired upper and lower quantile thresholds, as well as a number of standard deviations away from the mean to use for the calculations.

3. User sets the string value(s) in the DDI exports that should be considered as NA values.

4. The remainder of the script loops through the file list, returning the values listed below and binding them together by row into a single Excel spreadsheet that is saved to the User's working directory.

5. For each habitat included in the User's working directory a PDF report will be created in the "output" folder using *IQ_Report.Rmd*, which provides an overview of questionable / flagged values.

6. In addition to the PDF reports, each habitat will provide a .txt data output file in the "output/data" folder containing questionable values.

Summary

The following quantile thresholds are used for flagging "questionable" values:

- Lower quantile: **0.001**
- Upper quantile: **0.999**

Potential Included Parameters:

- Percent Cover - Species Composition
- Stem Density
- Total/Canopy Percent Cover - Species Composition

The data file used for the analysis: **All_CW_Parameters-2023-Jun-05.txt**

Table 1: Indicator Quantile Overview

parameter	q_low	q_high	mean	n_tot	n_q_low	n_q_high	%_flagg
Percent Cover - Species Composition	0	100.00	16.87	18681	0	0	0.
Stem Density	0	2446.56	76.26	3046	0	4	0.
Total/Canopy Percent Cover - Species Composition	0	103.22	43.51	150	0	0	0.

q_low: Value corresponding to the qval_low quantile for the parameter in the DDI export.

q_high: Value corresponding to the qval_high quantile for the parameter in the DDI export.

mean: Mean value for the parameter in the DDI export.

n_tot: Total number of records in the DDI export for the parameter.

n_q_low: Number of records in the DDI export that are below q_low for the parameter.

n_q_high: Number of records in the DDI export that are above q_high for the parameter.

%_flagged: Proportion of total records in the DDI export for the parameter which have been flagged as above q_high, or below q_low.

There are no Low Quantile flagged data entries for Coastal Wetlands

High Quantile

Stem Density

Table 2: Flagged Values - High Indicator Quantile: **2446.56**

RowID	ProgramID	ProgramLocationID	SampleDate	ResultValue
15533	4017	06A06	2015-10-21	2800
21321	4017	06C12	2015-10-21	2448
12299	4017	06C13	2015-10-21	2640
11857	4017	06C16	2015-10-21	2624

Programs containing flagged data:

4017 - Guana Tolomato Matanzas NERR Emergent Intertidal Vegetation Monitoring