Submerged Aquatic Vegetation 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:

- Braun Blanquet Score
- Modified Braun Blanquet Score
- Percent Cover
- Percent Occurrence
- Presence/Absence

The data file used for the analysis: **All_SAV_Parameters-2023-Oct-12.txt**

Table 1: Indicator Quantile Overview

parameter	sub_param	q_low	q_high	mean	n_tot	n_q_low	n_q_high	%_flagged
Braun Blanquet Score		0	5	0.32	1196087	1	0	0.00
Modified Braun Blanquet Score		0	5	2.19	30211	0	1	0.00
Percent Cover	Total	0	100	16.02	58529	0	43	0.07
Percent Cover	Species	0	100	2.48	440212	0	1	0.00
Percent Occurrence	Total	0	100	37.96	55987	0	2	0.00
Percent Occurrence	Species	0	100	7.59	542084	0	0	0.00
Presence/Absence		0	1	0.19	1356796	0	0	0.00

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.

Low Quantile

Braun Blanquet Score

Table 2: Flagged Values - Low Indicator Quantile: **0**

RowID	ProgramID	ProgramLocationID	SampleDate	CommonIdentifier	ResultValue
4726550	296	276	2020-12-13	Thalassia testudinum	-0.69

Programs containing flagged data:

296 - Florida Keys National Marine Sanctuary Seagrass Monitoring Project

High Quantile

Modified Braun Blanquet Score

Table 3: Flagged Values - High Indicator Quantile: **5**

RowID	ProgramID	ProgramLocationID	SampleDate	CommonIdentifier	ResultValue
4669136	559	KB06-2016	2016-06-30	Laurencia spp.	5.3

Programs containing flagged data:

559 - Northern Big Bend Seagrass Monitoring

Percent Cover - Species

Table 4: Flagged Values - High Indicator Quantile: **100**

RowID	ProgramID	ProgramLocationID	SampleDate	CommonIdentifier	ResultValue
4737539	564	W1-27-16	2016-09-30	Drift algae	590

Programs containing flagged data:

564 - Western Pinellas County Seagrass Monitoring

Percent Cover - Total

Table 5: Flagged Values - High Indicator Quantile: **100**

RowID	ProgramID	ProgramLocationID	SampleDate	CommonIdentifier	ResultValue
4742956	564	W1-06-12	2012-09-26	Total_SAV	105
4743445	564	W1-12-11	2011-10-05	Total_SAV	105
4741165	564	W1-16-17	2017-10-06	Total_SAV	101
4741893	564	W1-17-11	2011-09-30	Total_SAV	105
4741888	564	W1-18-14	2014-10-10	Total_SAV	101
4743589	564	W1-22-19	2019-11-10	Total_SAV	104
4743590	564	W1-22-19	2019-11-10	Total_SAV	150
4741145	564	W1-23-11	2011-10-03	Total_SAV	101
4741131	564	W1-24-11	2011-10-05	Total_SAV	105
4742264	564	W1-25-15	2015-10-02	Total_SAV	105
4743778	564	W1-26-12	2012-09-26	Total_SAV	101
4742245	564	W1-27-11	2011-10-04	Total_SAV	101
4742255	564	W1-27-11	2011-10-04	Total_SAV	140
4743008	564	W1-28-12	2012-09-26	Total_SAV	105
4743015	564	W1-28-12	2012-09-26	Total_SAV	105
4741395	564	W1-28-19	2019-12-31	Total_SAV	101
4743339	564	W1-31-15	2015-10-07	Total_SAV	110
4743349	564	W1-31-15	2015-10-07	Total_SAV	110
4743350	564	W1-31-15	2015-10-07	Total_SAV	110
4742415	564	W1-31-19	2019-12-31	Total_SAV	180
4743332	564	W1-34-15	2015-10-02	Total_SAV	125
4741656	564	W1-35-19	2019-12-31	Total_SAV	105
4741657	564	W1-35-19	2019-12-31	Total_SAV	110
4741658	564	W1-35-19	2019-12-31	Total_SAV	105
4741664	564	W1-35-19	2019-12-31	Total_SAV	105
4741665	564	W1-35-19	2019-12-31	Total_SAV	105
4741667	564	W1-35-19	2019-12-31	Total_SAV	110
4742039	564	W1-40-19	2019-12-31	Total_SAV	120
4742179	564	W2-01-12	2012-09-14	Total_SAV	105
4742855	564	W2-04-11	2011-09-29	Total_SAV	120
4741526	564	W2-04-12	2012-09-13	Total_SAV	140
4741521	564	W2-04-15	2015-10-21	Total_SAV	120
4741383	564	W2-2-19	2019-03-10	Total_SAV	130
4741393	564	W2-2-19	2019-03-10	Total_SAV	110
4743170	564	W2-5-19	2019-12-31	Total_SAV	120
4744197	564	W2-5-19	2019-12-31	Total_SAV	180
4741387	564	W2-9-19	2019-03-10	Total_SAV	122
4743236	564	W2-ALT-01-17	2017-09-29	Total_SAV	105
4743229	564	W3-03-15	2015-10-05	Total_SAV	110
4743231	564	W3-03-15	2015-10-05	Total_SAV	106
4743232	564	W3-03-15	2015-10-05	Total_SAV	125
4743238	564	W3-03-15	2015-10-05	Total_SAV	115
4743240	564	W3-03-15	2015-10-05	Total_SAV	135

Programs containing flagged data:

564 - Western Pinellas County Seagrass Monitoring

Percent Occurrence - Total

Table 6: Flagged Values - High Indicator Quantile: **100**

RowID	ProgramID	ProgramLocationID	SampleDate	CommonIdentifier	ResultValue
6085887	3013	IRLSG045	2021-08-12 09:15:00	Total seagrass	139
6100083	3013	IRLSG045	2021-08-12 09:15:00	Total seagrass	161

Programs containing flagged data:

3013 - Seagrass (SJRWMD)