Submerged Aquatic Vegetation Indicator Quantile Report

SEACAR Analysis

Last compiled on 20 June, 2024

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**

The following parameters are being excluded from this analysis:

Included Indicators and Parameters and the files used in this analysis:

All_SAV_Parameters-2024-Jun-06.txt

Indicator: Percent Cover

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

Summary Tables

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.

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

Indicator: **Percent Cover**

Table 1: Indicator Quantile Overview

ParameterName	q low	q high	mean	n tot	n q low	n q high	pct flagged
Braun Blanquet Score	0	5	0.36	1122266	1	4	0.00
Modified Braun Blanquet Score	0	5	2.05	44910	0	1	0.00
Percent Cover	0	100	4.36	525454	0	44	0.01
Percent Cover	0	100	4.68	453107	0	43	0.01
Percent Occurrence	0	100	10.30	623477	0	2	0.00
Percent Occurrence	0	100	10.83	507954	0	2	0.00
Presence/Absence	0	1	0.24	1856262	0	0	0.00

Low Quantile

Indicator: Percent Cover

Braun Blanquet Score

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

RowID	ProgramID	ProgramLocationID	SampleDate	CommonIdentifier	ResultValue
6643	296	276	2020-12-13	Thalassia testudinum	-0.69

Programs containing flagged data:

296 - Florida Keys National Marine Sanctuary Seagrass Monitoring Project

There are no *Low* Quantile Flagged Values for Modified Braun Blanquet Score

There are no *Low* Quantile Flagged Values for Percent Cover

There are no *Low* Quantile Flagged Values for Percent Occurrence

There are no *Low* Quantile Flagged Values for Presence/Absence

There are no *Low* Quantile Flagged Values for Shoot Count

High Quantile

Indicator: Percent Cover

Braun Blanquet Score

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

RowID	ProgramID	ProgramLocationID	SampleDate	CommonIdentifier	ResultValue
3202925	4018	13M	2018-07-19	Anadyomene stellata	15
3205960	4018	13M	2018-07-19	Anadyomene stellata	7
4008403	4018	13M	2018-07-19	Anadyomene stellata	10
3206197	4018	9P	2018-07-23	Anadyomene stellata	6

Programs containing flagged data:

4018 - Miami-Dade County DERM Benthic Habitat Monitoring Program

Modified Braun Blanquet Score

Table 4: Flagged Values - High Indicator Quantile: **5**

RowID	ProgramID	ProgramLocationID	SampleDate	CommonIdentifier	ResultValue
610214	559	KB06-2016	2016-06-30	Laurencia spp.	5.3

Programs containing flagged data:

559 - Northern Big Bend Seagrass Monitoring

Percent Cover

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

RowID	ProgramID	ProgramLocationID	SampleDate	CommonIdentifier	ResultValue
539368	564	W1-06-12	2012-09-26	Total SAV	105
536825	564	W1-12-11	2011-10-05	Total SAV	105
599188	564	W1-16-17	2017-10-06	Total SAV	101
593716	564	W1-17-11	2011-09-30	Total SAV	105
563966	564	W1-18-14	2014-10-10	Total SAV	101
567292	564	W1-22-19	2019-11-10	Total SAV	150
573474	564	W1-22-19	2019-11-10	Total SAV	104
541195	564	W1-23-11	2011-10-03	Total SAV	101
557902	564	W1-24-11	2011-10-05	Total SAV	105
564841	564	W1-25-15	2015-10-02	Total SAV	105
548561	564	W1-26-12	2012-09-26	Total SAV	101
541582	564	W1-27-11	2011-10-04	Total SAV	101
541867	564	W1-27-11	2011-10-04	Total SAV	140
542915	564	W1-28-12	2012-09-26	Total SAV	105
562629	564	W1-28-12	2012-09-26	Total SAV	105
567348	564	W1-28-19	2019-12-31	Total SAV	101
555589	564	W1-31-15	2015-10-07	Total SAV	110
592978	564	W1-31-15	2015-10-07	Total SAV	110
593062	564	W1-31-15	2015-10-07	Total SAV	110
564906	564	W1-31-19	2019-12-31	Total SAV	180
558609	564	W1-34-15	2015-10-02	Total SAV	125
570635	564	W1-35-19	2019-12-31	Total SAV	105
571915	564	W1-35-19	2019-12-31	Total SAV	105
577602	564	W1-35-19	2019-12-31	Total SAV	110
584430	564	W1-35-19	2019-12-31	Total SAV	105
586636	564	W1-35-19	2019-12-31	Total SAV	105
587110	564	W1-35-19	2019-12-31	Total SAV	110
594214	564	W1-40-19	2019-12-31	Total SAV	120
548722	564	W2-01-12	2012-09-14	Total SAV	105
539286	564	W2-04-11	2011-09-29	Total SAV	120
549450	564	W2-04-12	2012-09-13	Total SAV	140
555372	564	W2-04-15	2015-10-21	Total SAV	120
570411	564	W2-2-19	2019-03-10	Total SAV	110
596207	564	W2-2-19	2019-03-10	Total SAV	130
574147	564	W2-5-19	2019-12-31	Total SAV	180
580175	564	W2-5-19	2019-12-31	Total SAV	120
570410	564	W2-9-19	2019-03-10	Total SAV	122
594850	564	W2-ALT-01-17	2017-09-29	Total SAV	105
539053	564	W3-03-15	2015-10-05	Total SAV	135
541578	564	W3-03-15	2015-10-05	Total SAV	110
553719	564	W3-03-15	2015-10-05	Total SAV	106
555938	564	W3-03-15	2015-10-05	Total SAV	115
561052	564	W3-03-15	2015-10-05	Total SAV	125

Programs containing flagged data:

564 - Western Pinellas County Seagrass Monitoring

Percent Occurrence

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

RowID	ProgramID	ProgramLocationID	SampleDate	CommonIdentifier	ResultValue
2257094	3013	IRLSG045	2021-08-12 09:15:00	Total seagrass	161
2257880	3013	IRLSG045	2021-08-12 09:15:00	Total seagrass	139

Programs containing flagged data:

3013 - Seagrass (SJRWMD)

There are no *High* Quantile Flagged Values for Presence/Absence

There are no *High* Quantile Flagged Values for Shoot Count

QAQC Quantile Flag Check

- n_high is the amount of data above the quantile value.
- n_high_flagged is the amount of data above the quantile value AND containing proper SEACAR-QAQCFlag of 17Q (ResultValue above quantile value).
- n_low is the amount of data below the quantile value.
- n_low_flagged is the amount of data below the quantile value AND containing proper SEACAR-QAQCFlag of 16Q (ResultValue below quantile value).
- If everything is in order, these values should be the same. Any discrepancies therein should be fastidiously noted.

ParameterName	n high	n high flagged	n low	n low flagged
Braun Blanquet Score	4	0	1	1
Modified Braun Blanquet Score	1	1	0	0
Percent Cover	43	43	0	0
Percent Occurrence	2	2	0	0

Entries where ResultValue is above or below quantile, but expected SEACAR_QAQCFlagCode is not being applied

Table 7: SEACAR QAQC Flag Code discrepancies

RowID	ProgramID	ParameterName	ResultValue	SEACAR_QAQCFlagCode	q_subset
3202925	4018	Braun Blanquet Score	15	7Q/15Q	high
3205960	4018	Braun Blanquet Score	7	7Q/15Q	high
3206197	4018	Braun Blanquet Score	6	7Q/15Q	high
4008403	4018	Braun Blanquet Score	10	15Q/7Q	high

- 7Q - No defined thresholds for this parameter/15Q - Unexpected value: outside expected parameter values
- 15Q - Unexpected value: outside expected parameter values/7Q - No defined thresholds for this parameter