

Coastal Wetlands Indicator Quantile Report

SEACAR Analysis

Last compiled on 08 February, 2024

Contents

Overview	2
Purpose	2
Relevant file locations	2
Process steps	2
Summary	3
Summary Tables	4
Low Quantile	5
Indicator: Species Compositon	5
Stem Density	5
High Quantile	7
Indicator: Species Compositon	7
Stem Density	7
QAQC Quantile Flag Check	8

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:

- *Standard Length*

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

All_CW_Parameters-2024-Jan-10.txt

Indicator: Species Compositon

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

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.

%_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: **Species Compositon**

Table 1: Indicator Quantile Overview

parameter	q_low	q_high	mean	n_tot	n_q_low	n_q_high	%_flagged
Percent Cover	0.10	100.0	29.96	3977	0	0	0.00
Stem Density	0.01	5262.4	833.24	897	1	1	0.22

Low Quantile

Indicator: Species Compositon

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

Stem Density

Table 2: Flagged Values - Low Indicator Quantile: **0.01009**

RowID	ProgramID	ProgramLocationID	SampleDate	ResultValue
28283	906	Fort DeSoto	2016-11-15	0.01

Programs containing flagged data:

906 - Tampa Bay Critical Coastal Habitat Assessment

There are no *Low* Quantile Flagged Values for Total/Canopy Percent Cover

High Quantile

Indicator: Species Compositon

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

Stem Density

Table 3: Flagged Values - High Indicator Quantile: **5262.4**

RowID	ProgramID	ProgramLocationID	SampleDate	ResultValue
19788	5009	PC2-2	2014-11-04	5800

Programs containing flagged data:

5009 - Apalachicola Emergent Marsh Vegetation Monitoring

There are no *High* Quantile Flagged Values for Total/Canopy Percent Cover

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
Stem Density	1	1	1	0

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

Table 4: SEACAR QAQC Flag Code discrepancies

RowID	ProgramID	ParameterName	ResultValue	SEACAR_QAQCFlagCode	q_subset
28283	906	Stem Density	0.0099	19Q	low

- 19Q - Low threshold defined only/ not below