MassBays: Initial data exploration

emily shumchenia

4/30/2018

Embayment clusters

There are 4 clusters of embayments in MassBays, defined by the Northeastern assessment. Below they are described by which stressor and habitat metrics were highest.

Cluster 1
- highest saltmarsh extent
- highest tidal flat area

Cluster 2
- highest hardened shoreline
- highest seagrass extent

Cluster 3
- highest saltmarsh shoreline
- highest impairment for nutrients

Cluster 4
- highest high-intensity land use
- highest population density
- highest tidal restriction

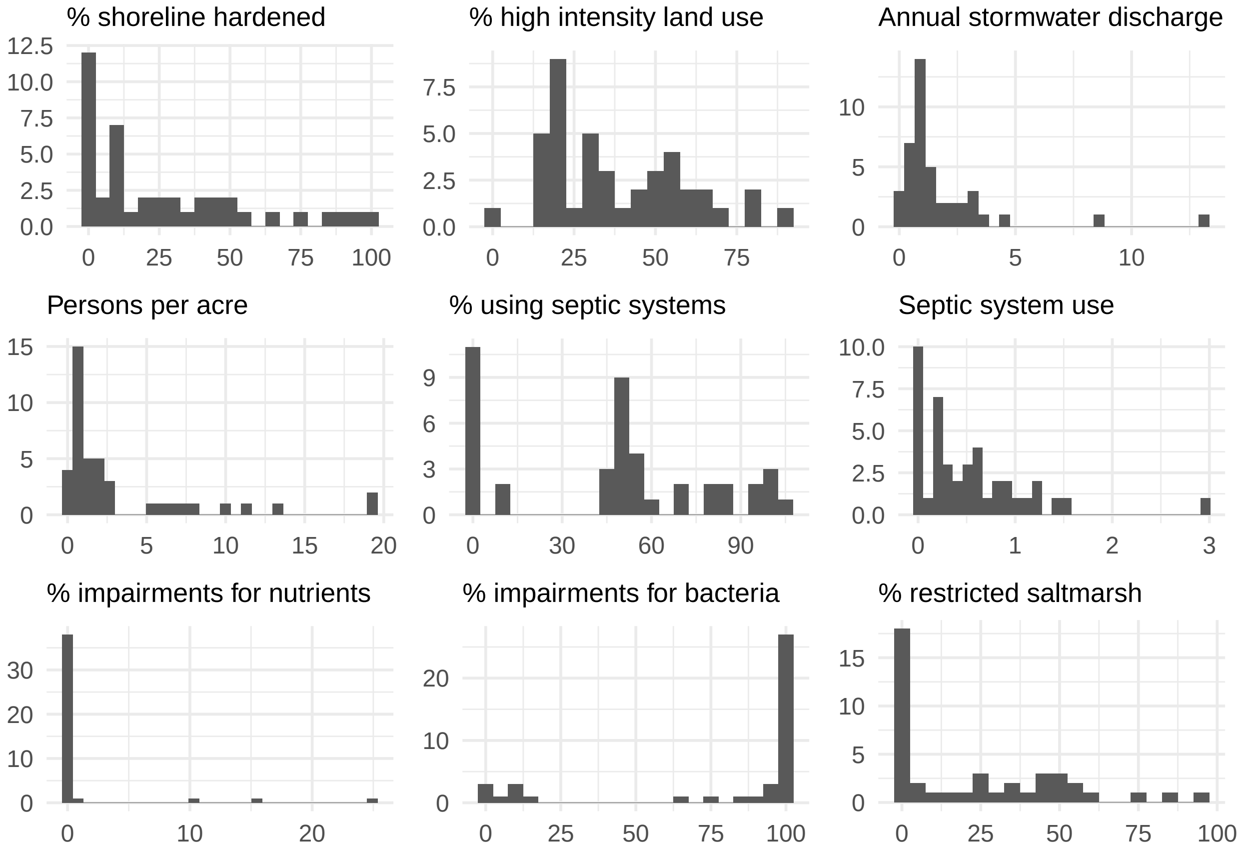
Stressor Data

Nine stressors variables were incorporated into the Northeastern assessment.

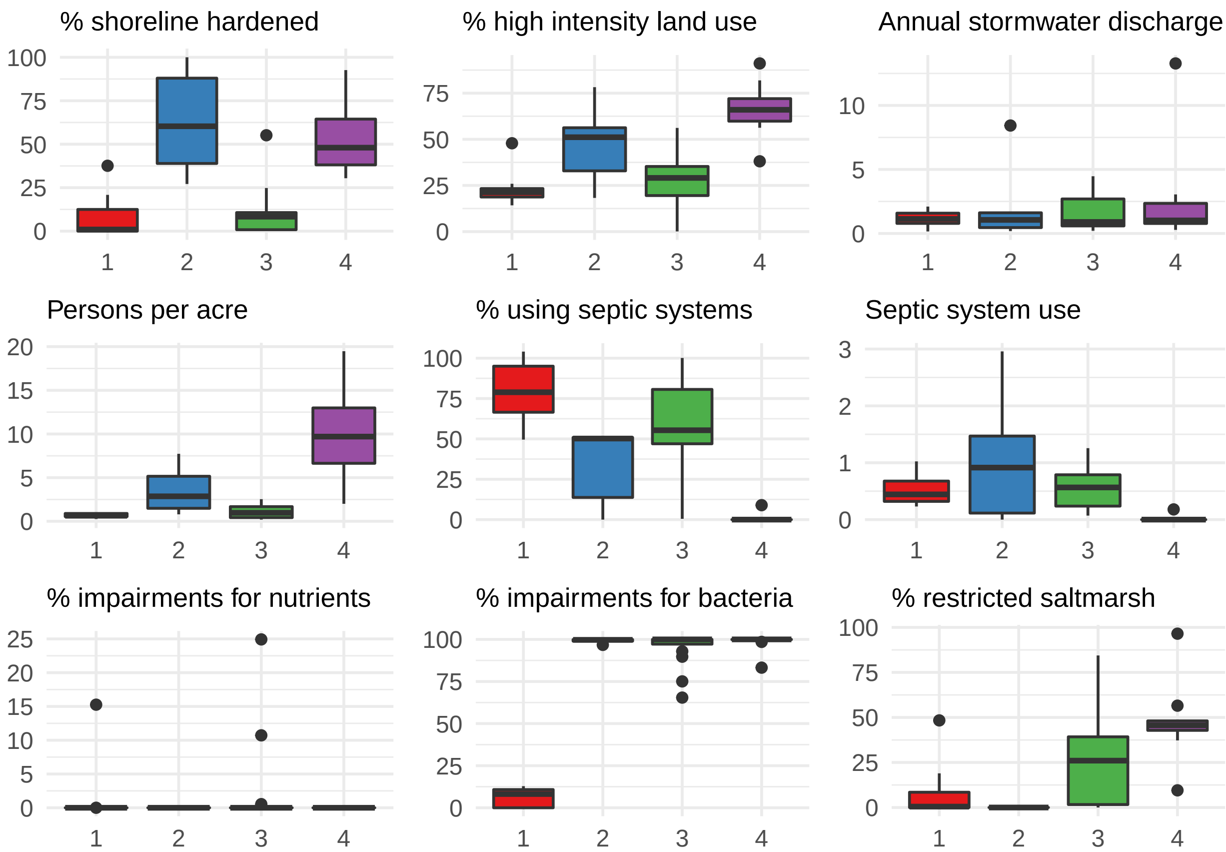
Table 1: Stressor variables

Indicator	Metric
% shoreline hardened	Linear kilometer of hardened shoreline divided by the sum of shoreline that was and that co
% high intensity land use	Acres high intensity land use (residential, commercial, industrial, agricultural, and transport
Annual stormwater discharge	Acre-ft stormwater / acres open water / year
Persons per acre	Number of people divided by the total acres in the embayment
% using septic systems	Number of people using septic divided by the total population
Septic system use	Number of people using septic systems divided by embayment land acreage
% impairments for nutrients	Acres of impaired water divided by the acres of open water
% impairments for bacteria	Acres of impaired water divided by the acres of open water
% restricted saltmarsh	Acres of salt marsh that are tidally restricted divided by the total acres of salt marsh

To understand the distribution of values for each stressor throughout MassBays, we can look at the histograms below.



Now we want to know how the embayment clusters differ with respect to these stressors. We can look at box plots showing the range of values for each stressor in each embayment cluster 1 - 4.



Resource Data (Habitat mosaic components)

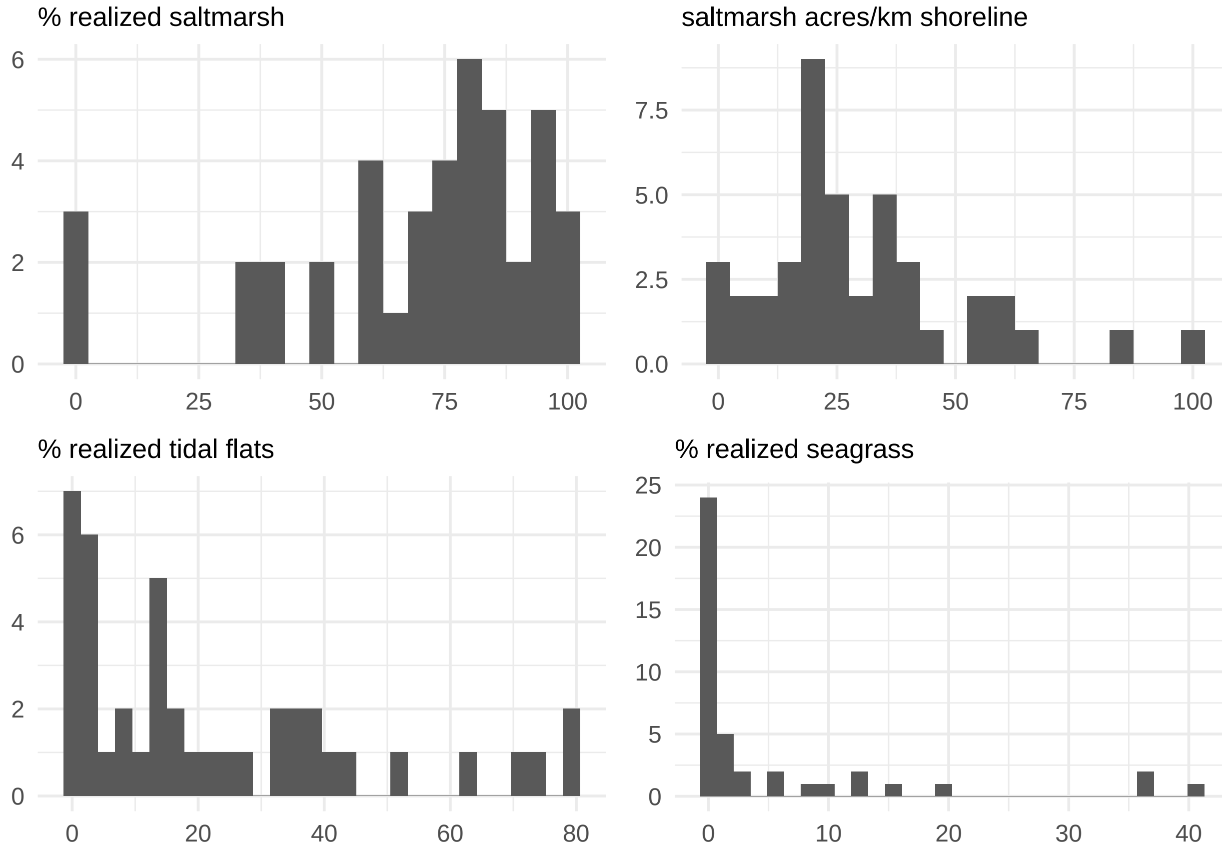
Four habitat metrics were examined by Northeastern: two for saltmarsh, one for seagrass, and one for tidal flats. Three of them are percentages of habitat existing with respect to the habitat available. For example, "% realized seagrass" is a measure of the acreage of seagrass that exists divided by the total area with depths suitable for eelgrass growth. In order to be comparable with the others, we rescaled saltmarsh acres on a scale from 0-100.

```
## Warning in read.table(file = file, header = header, sep = sep, quote
## = quote, : incomplete final line found by readTableHeader on 'data/
## resources_definitions.csv'
```

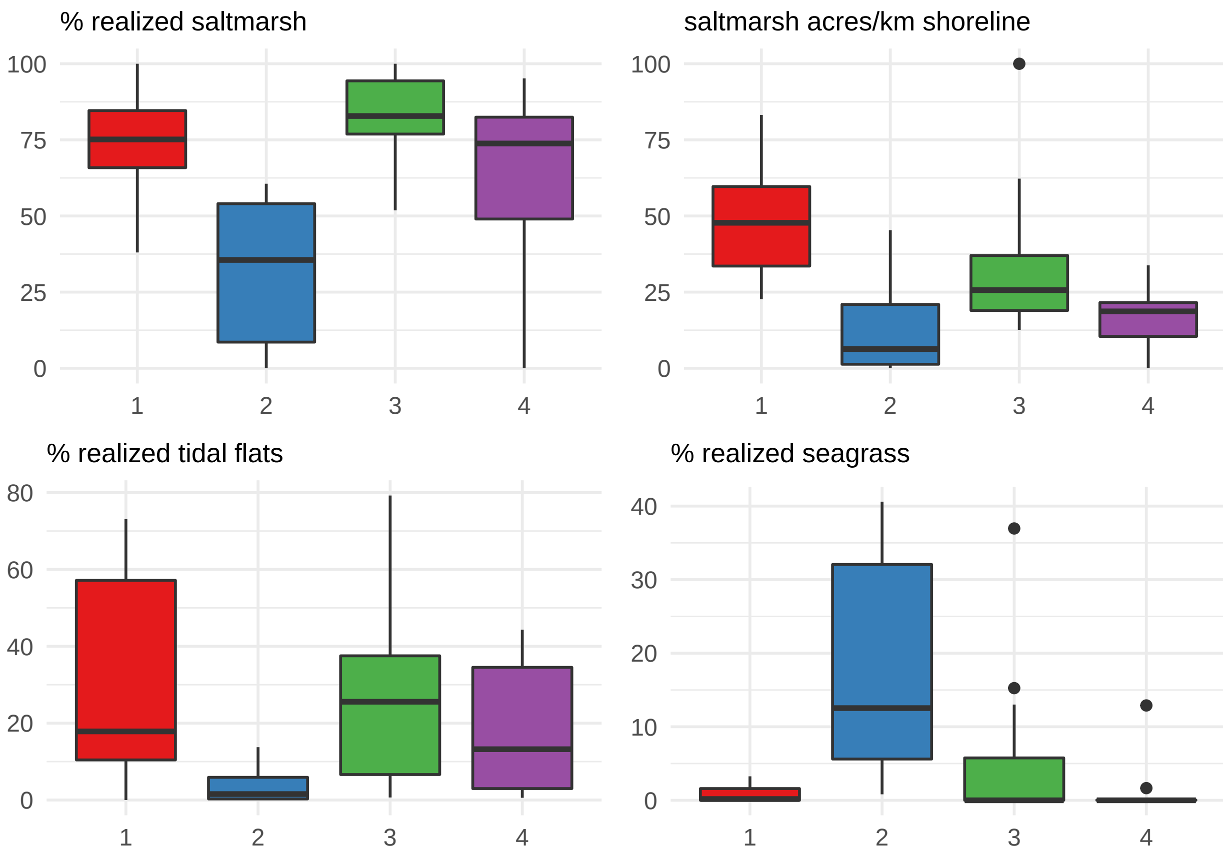
Table 2: Resource variables

Indicator	Metric
% realized saltmarsh	Linear kilometers of shoreline with salt marsh present divided by linear kilometers of shoreli
saltmarsh acres/km shoreline	Acres of salt marsh present in an embayment divided by the linear kilometers of shoreline w
% realized tidal flats	Acres of tidal flat divided by the acres of open water that are less than 5 meters in depth
% realized seagrass	Acres of seagrass divided by the acres of open water that are less than 10 meters in depth

The distribution of values for these metrics throughout all MassBays embayments is shown below.



Now we can also look at the amount of each habitat in each embayment cluster 1 - 4:

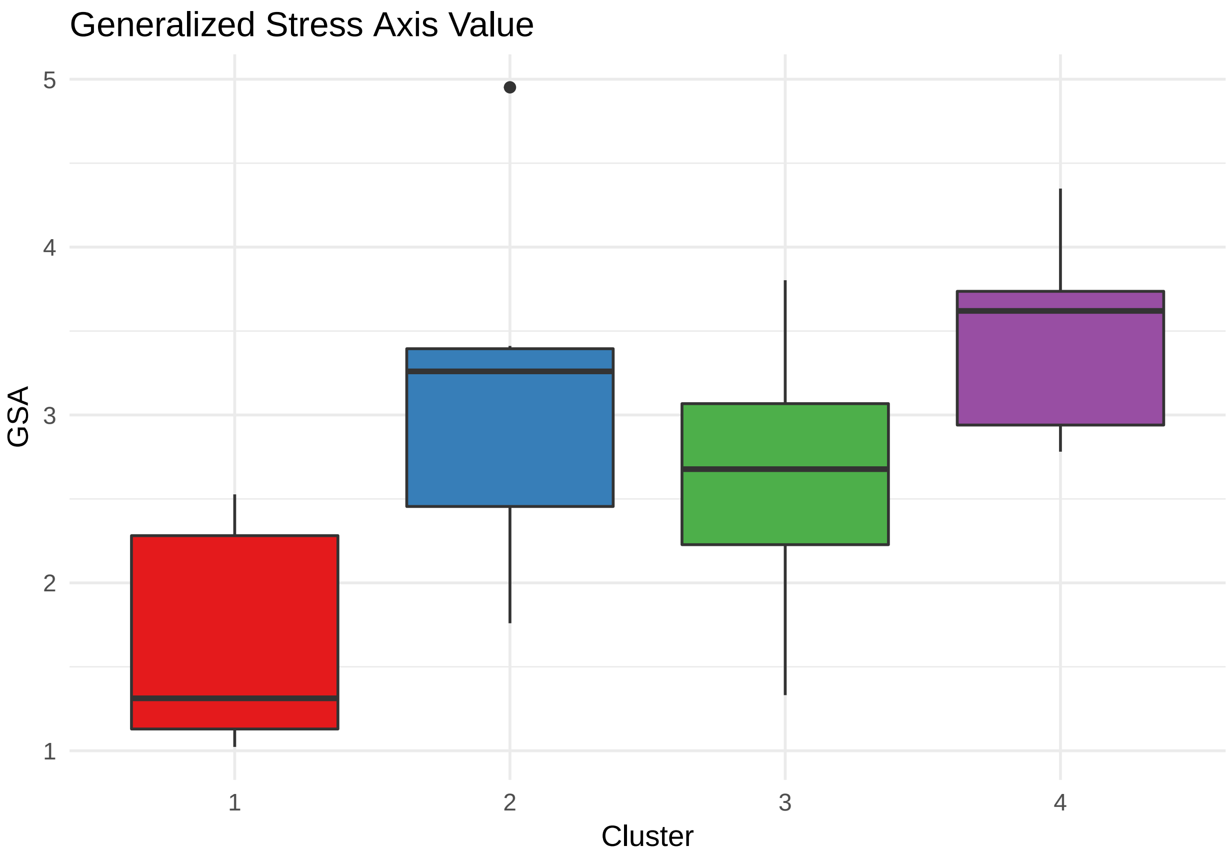


Apply these data to a BCG framework

To apply these data to a BCG framework, we'll plot the habitat data for each embayment in MassBays against a Generalized Stress Axis (GSA). This will also allow us to evaluate stressor-habitat characteristics in each embayment and cluster. The GSA is not meant to be used to examine causality between stressors and habitats; it is meant to depict the cumulative stress gradient.

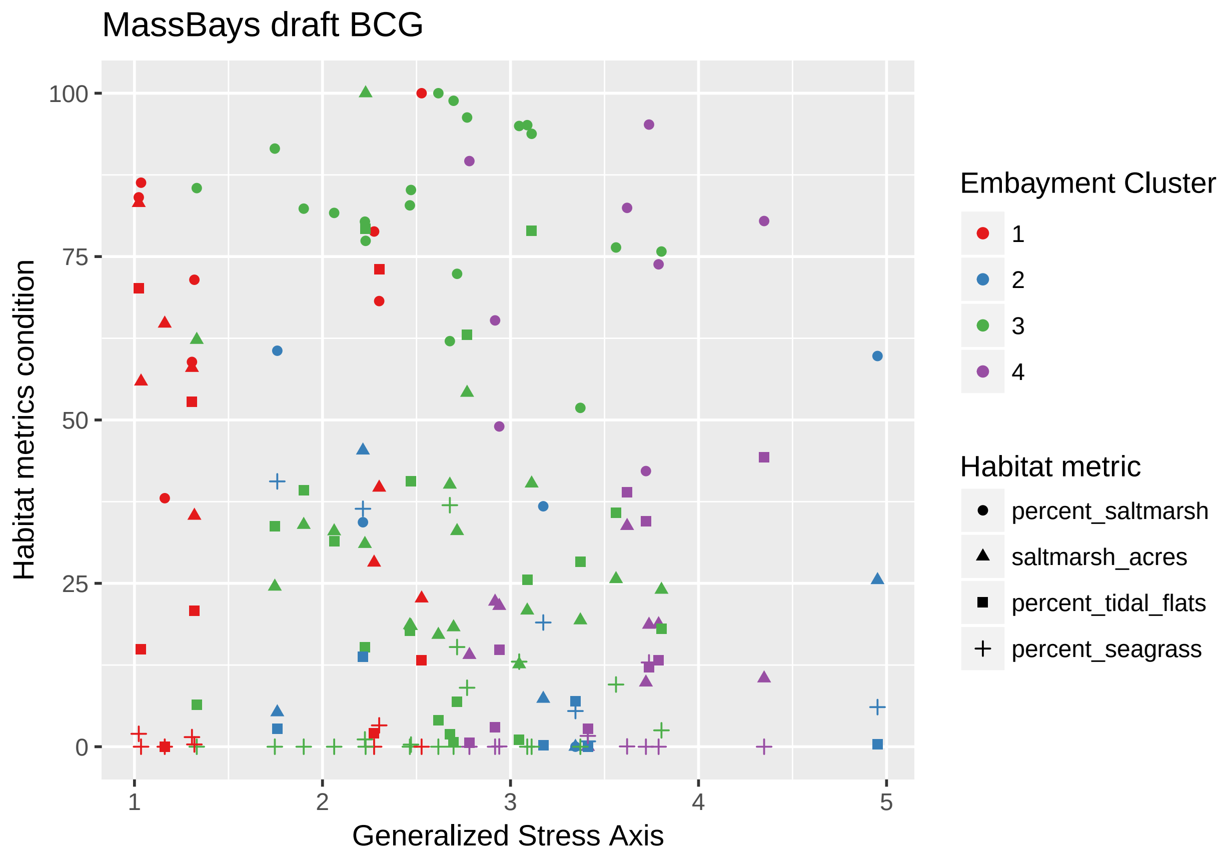
Developing a draft GSA

We'll use a simple formula to develop a draft GSA for all MassBays embayments together. First, we'll scale each stressor from 0-1 and then sum those values for each embayment in MassBays. Here are the values for a GSA in each cluster:



Assembling a draft "BCG-like" plot for all of MassBays

Now we want to plot the values of the habitat metrics in each MassBays embayment against the GSA. The resulting plot is very busy because there are 42 embayments and 4 habitat metrics in each. But, this plot could be viewed as "BCG-like". The y-axis is not broken up by BCG levels, but breaks between levels and definitions of levels could be defined.

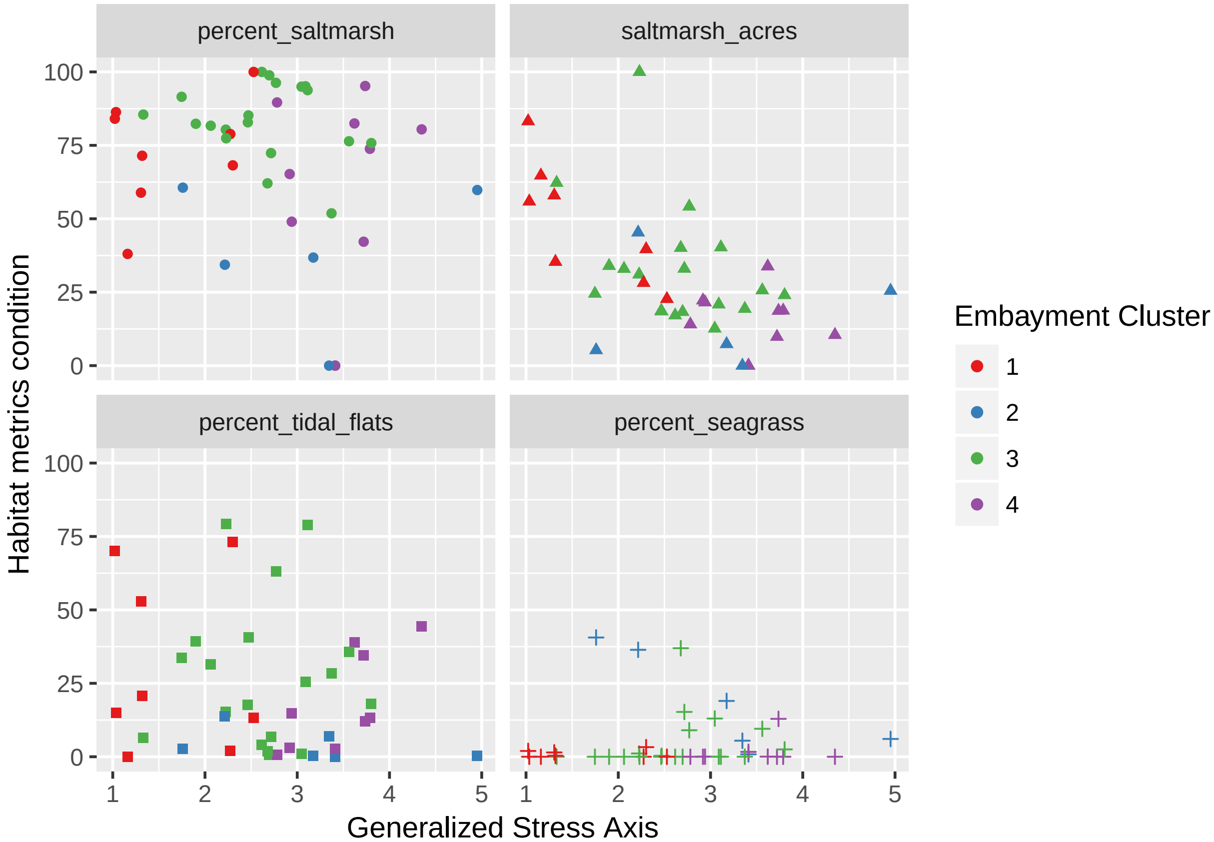


The plot can be simplified a number of ways. For example, each quadrant of the BCG represents a unique combination of stressor and habitat characteristics.

Individual "BCG-like" plots of each habitat type

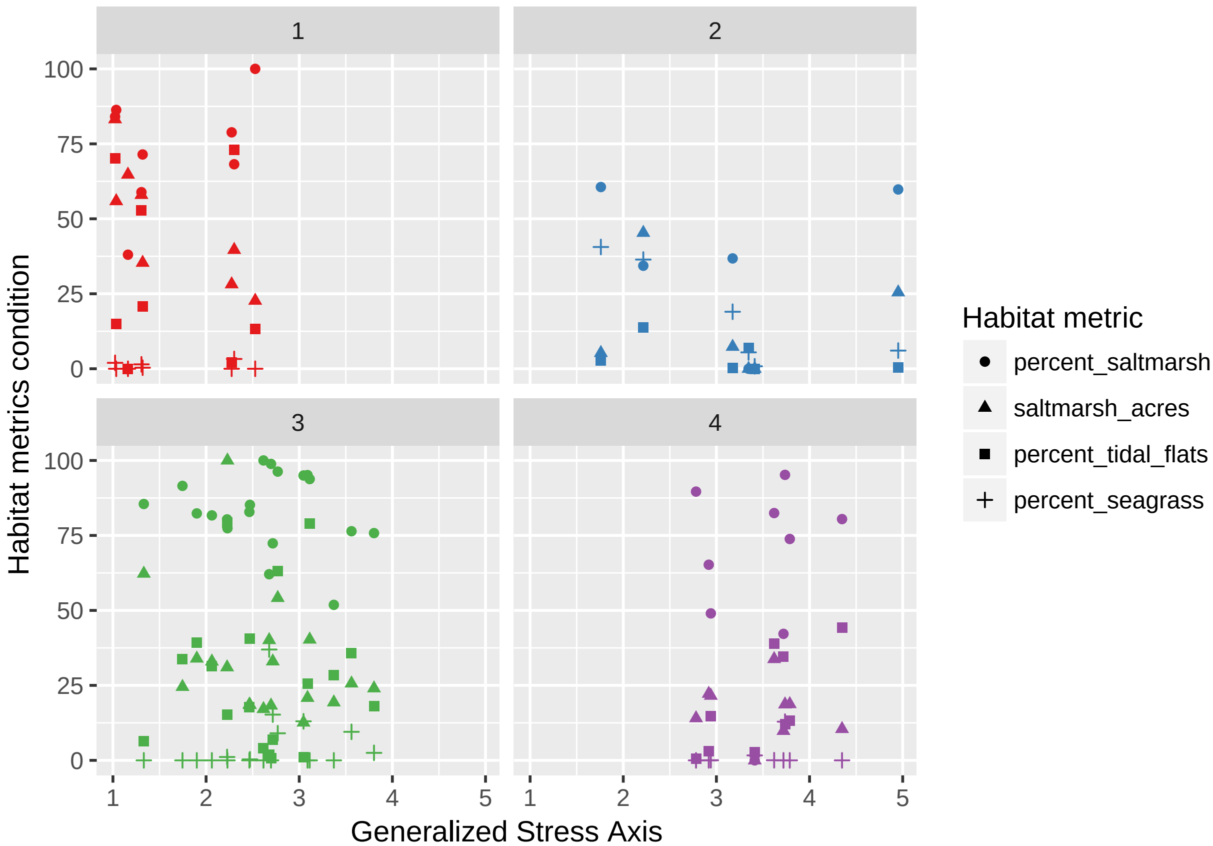
Another way to simplify is to look at a "BCG-like" plot for each habitat type separately. Again, the y-axis does not have BCG levels, but these could be developed.

We can see that Saltmarsh acres per km shoreline has the most linear relationship with the GSA.



Individual "BCG-like" plots for each embayment type

More likely, we would want to look at a BCG-like plot of each embayment cluster separately. They are plotted below and numbered 1-4. We can clearly see that the ranges of stressors and habitats differs among clusters. For example, Cluster 1 seems to be comprised of embayments with low exposure to stressors (the GSA is never greater than 3). On the other hand, despite high levels of exposure to stressors (the GSA is never less than about 2.5), Cluster 4 shows relatively high values for percent saltmarsh habitat.



Summary and Questions to Consider

What we've done so far is explore the stressors and habitat metrics within the bounds of MassBays as a whole. In doing so, we've shown that there are real differences among clusters that could be considered when assembling BCGs for each embayment.

In addition, these BCGs only consider the present range of stressor and habitat values captured in the EDA 2.0. A more complete BCG (or BCGs) may want to consider historical conditions (whether that be recent historical, pre-industrial, etc.)

The Northeastern assessment also identified which stressors are most predictive of habitat condition. Shoreline hardening and septic system use were able to predict percent saltmarsh; population density and shoreline hardening could predict saltmarsh acres per km shoreline. One option is to use only those stressors as inputs to a GSA for future draft BCGs. Another option could be to use the highest 3 stressors in each embayment cluster to develop custom GSAs for each cluster.

Questions to consider

1. How can additional stressor and/or habitat metrics be included (e.g., water quality, contaminants)?
2. Should restoration/conservation targets be based on present range of data in each cluster or should they be informed by additional (e.g., historical) data?
3. How can breaks between condition levels be defined? Is it possible to use the results of the Northeastern decision trees for this?
4. Are "custom-BCGs" for each embayment desired?