

Discriminant analysis

Discriminant analyses successfully differentiated the five geographic regions with the error rate of 6% (Table 11). Two Mid-Atlantic sites and one West Coast site were grouped with sites in the Northeast region, and one site located in the Northeast was grouped with sites in the Mid-Atlantic region. Attributes were most different between the West Coast and the Gulf of Mexico using the generalized square distance (Table 12). Attributes were most similar between the Northeast and the Mid-Atlantic. The Southeast and the Gulf of Mexico were also fairly similar when compared to other regions.

Differentiating attributes (Figure 23) were selected using three distinct stepwise discriminant analysis methods (stepwise, forward selection, and backward elimination, (Table 13). Stepwise and forward selection kept the same ten attributes in the model: warm water temperature ($\geq 25^{\circ}\text{C}$), cold water temperature ($\leq 10^{\circ}\text{C}$), $\text{pH} < 7$, $\text{pH} > 8$, clay, permeability, turbidity, precipitation, salinity range, and mean range in depth. Backward elimination was in general agreement with these two methods.

Table 11. Results of discriminant analysis of data from 51 NERR sites.

Error rates for Region						
	West Coast	Northeast	Mid-Atl.	Southeast	G. of Mex.	Total
Rate	0.0000	0.2143	0.1000	0.0000	0.0000	0.0629
Priors	0.2000	0.2000	0.2000	0.2000	0.2000	

Discriminant functions for five geographic regions.					
Attribute	West Coast	Northeast	Mid-Atlantic	Southeast	Gulf of Mexico
Constant	-10.74175	-9.53057	-3.60198	-18.65224	-21.07932
Watershed size	0.56332	-1.27959	-0.89539	1.77580	0.49671
Clay	1.50113	-3.89460	-1.60179	2.76667	3.98061
Agricultural Land	-1.51924	3.45665	3.39300	-4.62589	-2.98169
Forest Land	-1.22088	3.48660	3.20365	-3.61050	-4.64787
Urban Built-up	-0.90648	3.01824	2.25837	-3.76018	-2.59606
Wetland	-3.61723	1.25984	1.89716	1.06694	-1.58664
Shellfish	0.33203	1.45973	0.60462	-0.99846	-2.83115
Width at site	-0.40327	0.11864	0.71454	-0.91005	0.61813
Precipitation	-3.41186	-3.03165	-1.43486	3.60699	7.83432
Cold (<10oC)	-3.01713	4.41422	4.33346	-5.42449	-2.95964
Warm (>25oC)	-13.86439	-12.68825	-0.58623	18.15026	20.09128
Hypoxia	3.91375	2.10568	0.32169	-5.38441	-2.56996
Supersaturation	2.84222	3.28388	1.12462	-5.28405	-4.68603
Depth Range	2.05674	1.44691	0.31626	-1.46052	-4.19737
Salinity	-3.20406	-6.91758	-2.43461	10.09876	7.46359
Salinity Range	2.90932	-3.58488	-2.66735	3.20249	2.24909
pH<7	-1.01116	-5.09796	-0.82732	6.67713	3.28356
pH>8	-0.54280	-4.67166	-2.03469	6.30697	4.01550
Turbidity >25	-1.85558	-1.38186	0.45295	2.47769	1.22792
Perm	-1.88249	-2.52701	0.10663	1.23079	5.83271

Table 12. Dissimilarity between regions using generalized squared distance.

Region	West Coast	Northeast	Mid-Atlantic	Southeast	Gulf of Mexico
West Coast	0	24.98634	36.16749	86.52786	105.82605
Northeast	24.98634	0	13.15349	105.83015	107.75766
Mid-Atlantic	36.16749	13.15349	0	62.87010	59.25187
Southeast	86.52786	105.83015	62.87010	0	21.25898
Gulf of Mexico	105.82605	107.75766	59.25187	21.25898	0

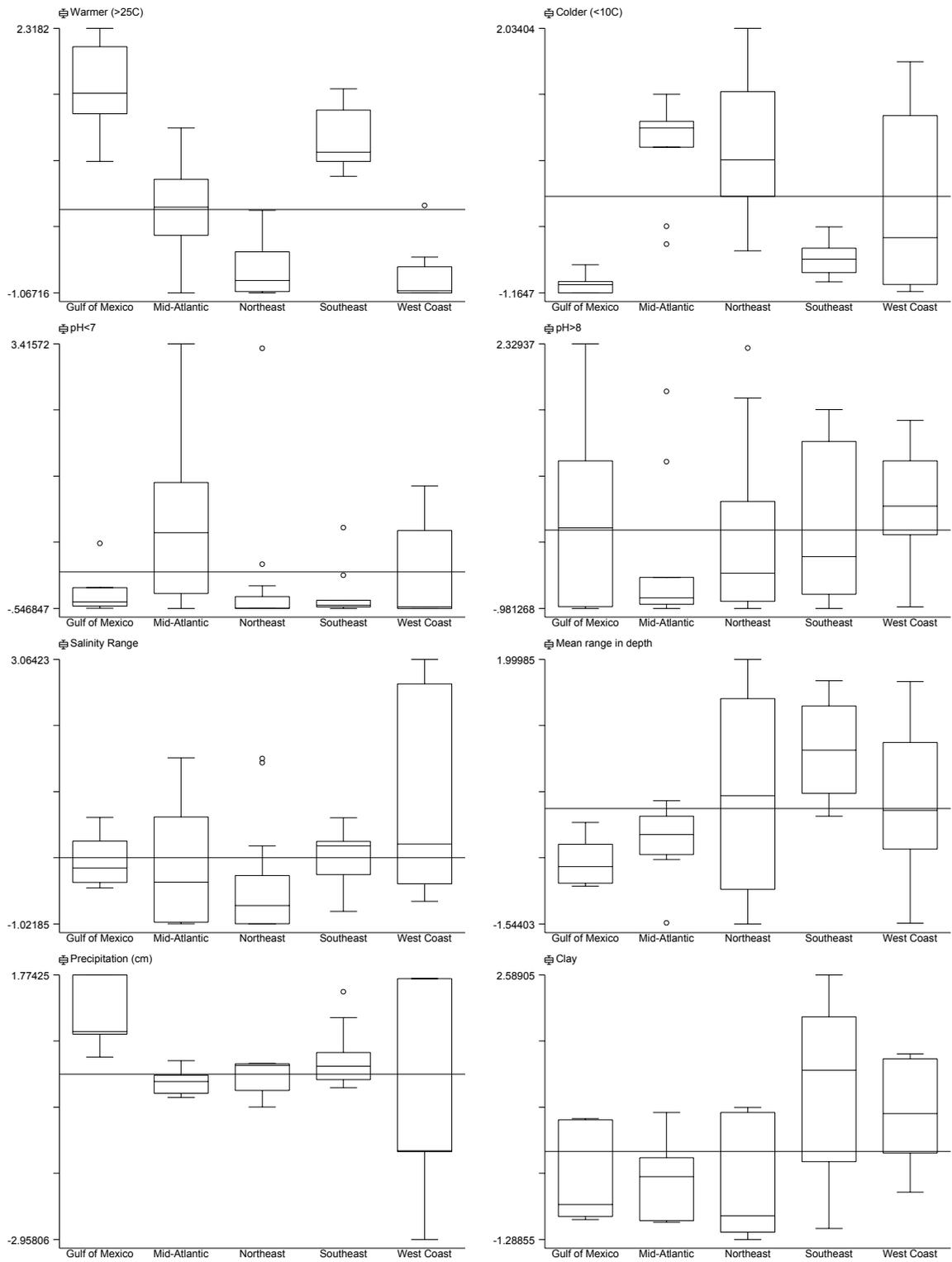


Figure 23. Box plots of differentiating attributes among five geographic regions. Each variable was standardized (with mean 0 and variance 1).

Table 13. Differentiating attributes for separating five geographic regions were selected using the stepwise discriminant analysis with three methods: stepwise selection, forward selection, and backward elimination (significance levels set at SAS default value of 0.15).

Summary of Stepwise selection

Step	Number		Entered	Removed	Partial		
	In				R-Square	F Value	Pr > F
1	1		warmer_25C		0.7991	45.74	<.0001
2	2		Clay		0.3899	7.19	0.0001
3	3		Perm		0.3973	7.25	0.0001
4	4		95_00_Precip_cm		0.3421	5.59	0.0010
5	5		SalRange		0.2907	4.30	0.0052
6	6		Colder_10C		0.3139	4.69	0.0033
7	7		Turb_25		0.2803	3.89	0.0092
8	8		DepRange		0.2274	2.87	0.0355
9	9		pH_8_0		0.1665	1.90	0.1309
10	10		pH_7_0		0.2009	2.32	0.0745

Summary of forward selection

Step	Number		Entered	R-Square	F Value	Pr > F
	In					
1	1		warmer_25C	0.7991	45.74	<.0001
2	2		Clay	0.3899	7.19	0.0001
3	3		Perm	0.3973	7.25	0.0001
4	4		95_00_Precip_cm	0.3421	5.59	0.0010
5	5		SalRange	0.2907	4.30	0.0052
6	6		Colder_10C	0.3139	4.69	0.0033
7	7		Turb_25	0.2803	3.89	0.0092
8	8		DepRange	0.2274	2.87	0.0355
9	9		pH_8_0	0.1665	1.90	0.1309
10	10		pH_7_0	0.2009	2.32	0.0745

Summary of backward elimination

Step	Number		Removed	R-Square	F Value	Pr > F
	In					
0	20					
1	19		width_atsite_m	0.0599	0.43	0.7858
2	18		Perm	0.0751	0.57	0.6879
3	17		Shellfish	0.0558	0.43	0.7868
4	16		watershed_HA	0.0892	0.73	0.5754
5	15		Forest_Land	0.1077	0.94	0.4566
6	14		Agricultural_Land	0.0494	0.42	0.7962
7	13		Urban_Built_up	0.0580	0.51	0.7304
8	12		Turb_25	0.1309	1.28	0.2971

Conclusions

In summary, eight distinguishing attributes were consistently important factors for all four methods including warm water temperature ($\geq 25^{\circ}\text{C}$), cold water temperature ($\leq 10^{\circ}\text{C}$), $\text{pH} < 7$, $\text{pH} > 8$, clay, precipitation, salinity range, and mean range in depth. As expected, temperature is the most distinguishing characteristic. Water temperature was very warm at sites in the Gulf of Mexico and fairly warm in the Southeast while water temperature in the Mid-Atlantic and the Northeast was cooler. Water was more acidic in the Mid-Atlantic when compared with other regions, where most sites were more oceanic and alkaline. Tidal dynamics (mean daily in depth, and salinity range) seemed greater in the Southeast and the West Coast than at sites in the Gulf of Mexico and the Mid-Atlantic. However, the variation of these attributes was large within the West Coast, so no clear-cut conclusion could be drawn. Precipitation was excessive in the Gulf of Mexico, and the proportion of clay was high in the Southeast and the West Coast.