Final Report

Project SR16-018: Continued Development of a Benthic Invertebrate Index for Barnegat Bay

Gary L. Taghon, Rutgers University, Project Manager
taghon@marine.rutgers.edu

Charlotte M. Fuller, Rutgers University, Co-Manager
fuller@marine.rutgers.edu

Rosemarie F. Petrecca, Rutgers University, Co-Manager and Quality Assurance Officer
petrecca@marine.rutgers.edu

Marc Ferko, NJDEP Quality Assurance Officer
Marc.Ferko@dep.state.nj.us

Bob Schuster, NJDEP Bureau of Marine Water Monitoring
Robert.Schuster@dep.state.nj.us

June 21, 2017

INTRODUCTION

The possibility of using benthic macroinvertebrates as indicators of water quality has been recognized for some time. For example, Hutchinson traced the origin and meaning of the term eutrophication (Hutchinson, 1973). In its earliest application to the trophic status and productivity of lakes, the species diversity of the bottom fauna, whether or not they could tolerate low oxygen conditions, and the organic content of sediment were all included in classification of lakes as oligotrophic or eutrophic. The potential has been difficult to realize in estuarine and coastal habitats, however, because macroinvertebrates respond to many environmental variables (e.g., temperature, salinity, dissolved oxygen) that can change quickly over space and time in these environments. Data collected over three years in Barnegat Bay, New Jersey, were used to explore if a straightforward relationship could be found between water quality and benthic macroinvertebrates (Taghon et al., 2015). The proportion of the total abundance consisting of those species most sensitive to nutrient pollution was used as the response variable. Exploratory data analysis identified summertime water total nitrogen concentration as the best, linearly correlated (negatively) variable, accounting for 84% of the variability in the proportion of sensitive species. Other potential variables (for example, salinity, chlorophyll-a, dissolved oxygen concentration) did not meet the assumptions of linear regression models.

This report includes a detailed analysis of the model. This report also includes the results of field sampling conducted in 2016. These samples were used to evaluate if the linear model could apply to "new" data that were not included in model development.

METHODS

Initial model development

Development of the model consisted of six steps: (1) compilation of water physical and chemical properties collected over time in Barnegat Bay-Little Egg Harbor estuary (BB–LEH) and analysis of seasonal patterns in water properties to identify times most likely to be stressful to benthic infauna; (2) selection of sites sampled for benthic infauna that were near sites where water properties were measured; (3) categorization of benthic infaunal species into ecological groups according to their sensitivity to environmental perturbations; (4) exploratory data analysis to identify possible patterns in water quality and the abundance of the most sensitive species of benthic infauna; (5) multiple linear regression analysis to identify the relationship between explanatory and dependent variables, and subsequent verification of the assumptions of linear regression; (6) validation of the regression model.

Step 1. Data on water temperature, salinity, total nitrogen concentration, total phosphorus concentration, chlorophyll *a* concentration, and dissolved oxygen concentration over the time period from 03/21/2011 to 12/16/2014 were downloaded from the EPA STORET Data Warehouse (http://www.epa.gov/storet/). If the data set contained multiple values for a variable for a given date, for example multiple samples collected throughout the day, the average value for that date was used. As expected, all were at levels likely to be most stressful during the months July–September (see Taghon et al. 2015 for details). The averages of all variables for those three-month periods in each year were calculated.

Step 2. Benthic infauna were sampled at 100 locations throughout BB–LEH in 2012, 2013, and 2014 (Taghon et al., 2013; Taghon et al., 2014; Taghon *et al.*, 2015). The same locations were sampled in July of each year. We chose the subset of those benthic infauna stations that were within a 2 km radius of locations where water property data were available (from *Step 1*). This cut-off distance resulted in using 59 of the 100 benthic stations for subsequent model development. We matched benthic stations that were sampled one year after water quality data were available, reasoning that environmental conditions, if stressful, would affect the abundance (due to survival, reproduction, and recruitment) of benthic infauna the following year. This resulted in 17 water property–benthic infauna groupings (Table 1 and Figure 1).

Step 3. Benthic infauna species were assigned to one of five Ecological Groups, based on their tolerance or response to organic enrichment: sensitive, indifferent, tolerant, second-order opportunists, and first-order opportunists. These Ecological Groups are described by Grall and Glémarec (1997):

> *"Group 1: Species very sensitive to organic enrichment and present in normal conditions.*
> *Group 2: Species indifferent to enrichment, always present in low densities with non-significant variations in time.*
> *Group 3: Species tolerant of excess organic matter enrichment. These species may occur in normal conditions but their populations are stimulated by organic enrichment.*
> *Group 4: Second-order opportunistic species. These are the small species with a short life cycle, adapted to a life in reduced sediment where they can proliferate.*
> *Group 5: First-order opportunistic species. These are the deposit feeders that proliferate in sediments reduced up to the surface."*

Rare species, defined here as any species for which fewer than 10 individuals were collected in all samples combined, in all three years, were omitted. Species were assigned to Ecological Groups (Table 2) using accepted published criteria (Borja, Mader, and Muxika, 2012; Gillett et al., 2015; Grall and Glémarec, 1997).

Step 4. Abundances per sample of all species in each Ecological Group were converted to proportional abundance, necessary to allow combining samples that had varying absolute abundances. Data for water properties and the proportion of Ecological Group 1 (EG1) for each of the 17 groups (Table 1) were entered into a spreadsheet (Table 2) for exploratory data analysis (Zuur, Ieno, and Elphick, 2010). Scatterplot matrices were constructed separately for each of the three water property–benthic infauna comparisons. Locally weighted regression (LOESS) was used to explore potential relationships between variables (Cleveland and Devlin, 1988). All statistical analyses were conducted using Statistix v10 (Analytical Software, Tallahassee, Florida).

The proportion of EG1 species in July 2012 was not correlated with water properties in the previous summer (Figure 2A). The proportion of EG1 species in 2013 showed general linear increases with salinity and decreases with total N from the previous summer, but with considerable scatter and a gap with no data at intermediate values of salinity and total N (Figure 2B). The proportion of EG1 species in 2013 showed a negative trend with chlorophyll concentration, but a positive trend with total P. In 2014 the most consistent trends were a decrease in the proportion of EG1 species with total N and total P (Figure 2C). The proportion peaked at intermediate values of chlorophyll and salinity.

The combined data set for all years showed generally linear relationships between the water properties and the proportion of EG1 species for chlorophyll concentration (negative), salinity (positive), temperature (negative), and total N concentration (negative) (Figure 2D). There was no consistent pattern with dissolved oxygen or total P.

Step 5. Stepwise linear regression of the combined data in Figure 2D resulted in a model with only total N as the independent variable (Table 3). Three key assumptions of linear regression were evaluated by examination of the residuals, the differences between actual values and values predicted from the regression model (Boldina and Beninger, 2016).

Assumption I. Independent variable is uncorrelated with the residuals. This assumption was tested by plotting the standardized residuals (each residual divided by the standard deviation of all residuals) against the total N concentration (Figure 3A). The residuals were distributed symmetrically about a value of zero; this assumption was met.

Assumption II. Homoscedasticity (homogeneity of variance); variance of the residuals is the same across the range of the independent variable. This assumption was tested by plotting the standardized residuals against the proportion of EG1 species predicted by the model (Figure 3B). The spread of the residuals increased with the predicted value, indicating that this assumption was not met.

Assumption III. Residuals are normally distributed. This assumption was tested by plotting the standardized residuals against their rankits (Figure 3C). Normally distributed residuals will lie on a straight diagonal line. There was slight departure from a straight line, but the residuals were considered normally distributed (Shapiro-Wilk statistic = 0.91, p = 0.094).

Step 6. The scatterplot for all three years of total N concentration and proportion of EG1 species shows the extent of spatial and temporal variation in these variables (Figure 4). For all data combined, the initial regression model was

$$\text{Proportion sensitive species} = 0.516\text{-}5.79\text{E-}04\times\text{Total N} \quad (1)$$

The regression coefficient was highly significant (t=-5.66, p<0.00001). Because of the high spatial and temporal variability, however, the coefficient of determination was low (adjusted R^2=0.214).

Final model development

The final model was based on reducing the spatial and temporal variability in the variables. This was done by calculating averages of the variables in each of the 17 groups used in the initial model (Table 4). The proportion of EG1 species showed consistent trends with only chlorophyll concentration, salinity, and total nitrogen concentration (Figure 5), therefore only these independent variables were entered into a stepwise linear regression. Chlorophyll concentration and salinity accounted for little of the variability in the proportion of sensitive species and were dropped from the final model (Table 5).

The final model met all three key assumptions of linear regression: independent variable (total N) was uncorrelated with residuals (Figure 6A), homogeneity of variance of residuals (Figure 6B), and normal distribution of residuals (Figure 6C). The final model was

$$\text{Proportion sensitive species} = 0.553\text{-}6.54\text{E-}04\times\text{Total N} \quad (2)$$

The regression coefficient was highly significant (t=-9.38, p<0.00001). The average proportion of sensitive species decreased as total N increased, with total N accounting for 84% of the variability (adjusted r^2 = 0.845).

Sampling in 2016

Benthic sampling locations were based on stations where water quality data from 2015 were available (data supplied by R. Schuster, NJDEP). Benthic samples at 14 stations were collected in July 2016 (Table 6). At each station, three 0.04-m^2 Ted Young Modified Van Veen grabs were taken. Two grab samples were immediately processed on board for invertebrate macrofauna analysis. Sediment was sieved over a 0.5-mm-mesh screen. The residue remaining on the screen was fixed in 3.7% formaldehyde solution in seawater, buffered with Borax and containing Rose Bengal to stain organisms.

The third grab sample was used for measurement of sediment properties. The top 2-cm layer of sediment was removed, transferred to a stainless steel bucket, and homogenized by stirring with a stainless steel spoon. Subsamples of the homogenized sediment were taken for elemental analysis (100 cm^3 of sediment transferred to a glass 250 mL jar with a Teflon-lined cap) and for grain size analysis (250 cm^3 of sediment transferred to a Whirl-Pak bag). These sediment samples were stored on ice following collection and during transport to the laboratory.

At the time of benthic sampling, surface and bottom water salinity, temperature, dissolved oxygen, and pH at each station were measured using a YSI Professional Plus meter.

Invertebrate samples were sorted and identified to the lowest practical taxonomic unit, usually species, by technicians at Cove Corporation (10200 Breeden Rd, Lusby, MD 20657). Validity of each species identified were checked against the continuously updated list in the World Registry of Marine Species (WoRMS) or the Integrated Taxonomic Information System (ITIS) to bring species names up to date with their current taxonomic status, thus facilitating future analyses of the benthic data, and allowing for cross study comparisons. Certain invertebrates that are unlikely to be fully sampled using the 0.5-mm screen proposed for this study (e.g. oligochaetes) were only identified to the class or genus level. For consistency with analyses of the benthic community in previous years (Taghon *et al.*, 2013; Taghon *et al.*, 2014; Taghon *et al.*, 2015), species for which fewer than 10 individuals were collected in all samples combined were excluded from analysis. Species were placed into ecological groups using the categories from Borja et al. (2012).

Sediment for elemental analysis was dried at 60°C and homogenized using an agate mortar and pestle. Any visible shell fragments, plant fragments, or obvious debris were removed prior to homogenization. Total carbon, nitrogen, and phosphorus of sediment were measured using standard methods (elemental analysis EPA Method 440.0 for total C and N (US EPA, 1992), colorimetric analysis of total phosphate for P (US EPA, 2010)). Two replicates per station were analyzed.

Sediment for grain size analysis was processed using methods described in detail in the EMAP-Estuaries Laboratory Methods Manual (US EPA, 1995). Sediment was wet-sieved through a 63µm-mesh sieve using 10% sodium hexametaphosphate solution in distilled water as dispersant to separate the silt and clay fraction (<63 µm) from the sand-sized fraction (>63 µm). The silt/clay suspension was transferred to a 1000-mL graduated cylinder, which was then fill to the mark with 10% sodium hexametaphosphate. The suspension was mixed thoroughly, then 20mL was withdrawn and transferred to a pre-weighed 50mL beaker. After drying at 60°C, the beaker and sediment was re-weighed to determine the mass of the silt/clay fraction, corrected for the dilution to the mass in the original sample. The sand-sized fraction was dried at 60°C, then separated with stacked graded sieves [63-125 µm fraction (very fine sand), 125-250 µm fraction (fine sand), 250-500 µm fraction (medium sand), 500-1000 µm fraction (coarse sand), and >1000 µm fraction (very coarse sand)] for 10 minutes on a sieve shaker. Each fraction was weighed. Grain size statistics were computed using the United States Geological Society software program GSSTAT (Poppe, Eliason, and Hastings, 2004).

RESULTS and DISCUSSION

The ranges of the environmental and biological data from the 2016 samples must not exceed the ranges of those parameters measured previously in order to allow a valid test of the model. In other words, extrapolation beyond the data used to develop the model is not permissible. Even environmental variables that were not used in development of the model should be within the ranges of those variables measured previously to minimize the possibility that an unusual event affected the dependent variable, in this case the proportion (relative abundance) of benthic invertebrates characterized as sensitive species. For example, water temperature had no effect in the model but unusually high or low water temperatures in 2016, relative to those in prior years, could have affected the benthic community structure.

Environmental data from 2016 were not anomalous, compared with previous years. Surface and bottom water dissolved oxygen concentration, pH, salinity, and temperature at the time of benthic sampling in 2016 (Table 6) were all within the ranges measured during benthic sampling in 2012–2014 (Figure 7). Sediment particle size distributions and sorting coefficients at the time of benthic sampling in 2016 (Table 7) were all within the ranges measured in 2012–2014 (Figure 8). Sediment total C and total P concentrations were slightly higher at a few stations in 2016 relative to prior years, while total N was not.

Benthic community structure in 2016 was similar to that in 2012–2014. Of the 184 taxa collected in 2012–2014, 152 were collected in 2016 (Table 8). The fewer number of taxa collected in 2016 was due to the fewer number of stations sampled then (14), relative to previous years (Table 1). In fact, when years are compared based on similar numbers of stations sampled, the species richness in

2016 was greater than that in prior years (Figure 9). Fifty-seven of all taxa collected in 2012–2014 were classified as ecologically sensitive, while 45 of these sensitive taxa were also collected in 2016.

The water total N concentrations at 14 stations sampled throughout 2015 were provided by R. Schuster, NJDEP. Average total N values over the months July–September were computed and used to calculate the predicted proportional abundance of all sensitive species at each station. Then, the actual proportions of sensitive species at co-located benthic stations were compared with those predicted by Equation 2 (Table 9). Because these data represent only one year, July–September 2015 for water quality matched with July 2016 for benthic invertebrates, it is necessary to put them into the perspective of annual variability in order to evaluate the goodness-of-fit of the model. The data from 2015 for water quality and 2016 for benthic invertebrates fell within the bounds of individual-year data used to develop the model (Figure 10A). More importantly, the residuals (differences between actual values and model predictions) were not different between the data used for the model and the "new" data (Figure 10B, Table 10).

Further sampling will provide an increasingly rigorous test of the model as the gap in time between the data used to develop the model and newly collected data widens.

References

Boldina, I. and Beninger, P.G., 2016. Strengthening statistical usage in marine ecology: Linear regression. *Journal of Experimental Marine Biology and Ecology*, 474, 81-91.

Borja, A.; Mader, J., and Muxika, I., 2012. Instructions for the use of the AMBI index software (Version 5.0). *Revista de Investigación Marina, AZTI-Tecnalia*, 19(3), 71-82.

Cleveland, W.S. and Devlin, S.J., 1988. Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83(403), 596-610.

Gillett, D.J.; Weisberg, S.B.; Grayson, T.; Hamilton, A.; Hansen, V.; Leppo, E.W.; Pelletier, M.C.; Borja, A.; Cadien, D.; Dauer, D.; Diaz, R.; Dutch, M.; Hyland, J.L.; Kellogg, M.; Larsen, P.F.; Levinton, J.S.; Llansó, R.; Lovell, L.L.; Montagna, P.A.; Pasko, D.; Phillips, C.A.; Rakocinski, C.; Ranasinghe, J.A.; Sanger, D.M.; Teixeira, H.; Dolah, R.F.V.; Velarde, R.G., and Welch, K.I., 2015. Effect of ecological group classification schemes on performance of the AMBI benthic index in US coastal waters. *Ecological Indicators*, 50, 99-107.

Grall, J. and Glémarec, M., 1997. Using biotic indices to estimate macrobenthic community perturbations in the Bay of Brest. *Estuarine, Coastal and Shelf Science*, 44, Supplement 1(0), 43-53.

Hutchinson, G.E., 1973. Marginalia: Eutrophication: The scientific background of a contemporary practical problem. *American Scientist*, 61(3), 269-279.

Poppe, L.J.; Eliason, A.H., and Hastings, M.E., 2004. A Visual Basic program to generate sediment grain-size statistics and to extrapolate particle distributions. *Computers & Geosciences*, 30(7), 791-795.

Taghon, G.L.; Grassle, J.P.; Fuller, C.M.; Petrecca, R.F., and Ramey, P., 2013. *Benthic Invertebrate Community Monitoring and Indicator Development for Barnegat Bay-Little Egg Harbor Estuary. Final Report, Project SR12-002. Department of Marine and Coastal Sciences, Rutgers University.*p.

Taghon, G.L.; Grassle, J.P.; Fuller, C.M.; Petrecca, R.F., and Ramey, P., 2014. *Benthic Invertebrate Community Monitoring and Indicator Development for Barnegat Bay-Little Egg Harbor Estuary: Year 2. Final Report, Project SR12-006. Department of Marine and Coastal Sciences, Rutgers University.*p.

Taghon, G.L.; Grassle, J.P.; Fuller, C.M.; Petrecca, R.F., and Ramey, P., 2015. *Benthic Invertebrate Community Monitoring and Indicator Development for Barnegat Bay-Little Egg Harbor Estuary: Year 3. Final Report, Project SR14-004. Department of Marine and Coastal Sciences, Rutgers University.*p.

US EPA, 1992. *Methods for the Determination of Chemical Substances in Marine and Estuarine Environmental Sample. United States Environmental Protection Agency, Office of Research and Development, Washington DC. EPA/600/R-92/121*p.

US EPA, 1995. *Environmental Monitoring and Assessment Program (EMAP): Laboratory Methods Manual - Estuaries, Volume 1: Biological and Physical Analyses. United States Environmental Protection Agency, Office of Research and Development, Narragansett, RI. EPA/620/R-95/008*p.

US EPA, 2010. *Sampling and Analytical Procedures for GLNPO's Open Lake Water Quality Survey of the Great Lakes. United States Environmental Protection Agency, Great Lakes National Program Office, Chicago, IL. EPA 905-R-05-001.*p.

Zuur, A.F.; Ieno, E.N., and Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1(1), 3-14.

Table 1. Water property and benthic infauna stations in the 17 groups based on similar locations in BB–LEH; see Figure 1 for location map. Years (in parentheses) of data used for model development.

Group #	Water quality stations	Benthic infauna stations
1	1826A (2013)	93, 95, 97, 98 (2014)
2	1818D/BB13 (2011–2012)	83, 84 (2012–2013)
3	1834A/BB12 (2011–2013)	79, 88, 91, 94, 96 (2012–2014)
4	BB11/BB11a (2011–2012)	74, 75, 85, 89 (2012–2013)
5	1707C (2013)	62, 68, 69, 73 (2014)
6	BB10 (2011–2012)	59 (2013), 60 (2012–2013)
7	1674B/BB09 (2011–2013)	59 (2012), 37, 39, 42, 51, 52, 56, 63 (2012–2014)
8	1691A/BB07/BB07a (2011–2013)	40, 43, 45, 48, 50 (2012–2014)
9	1661F (2013)	49, 53, 54, 58 (2014)
10	BB06 (2011–2013)	26, 34, 36 (2012–2014)
11	BB05 (2011)	21, 28, 29, 30, 33 (2012)
12	BB05a (2012)	16, 17, 22, 31, 32 (2013)
13	BB04 (2011)	20 (2012)
14	BB04a (2012–2013)	15, 24 (2013–2014)
15	BB03 (2011–2013)	1 (2012–2014)
16	BB02 (2011–2012)	2, 4, 6, 8 (2012–2013)
17	BB01 (2011–2013)	11 (2012–2014)

Table 2. Physical and chemical water properties for each station group (Table 1, Figure 1), average values for July–September for each year. Corresponding stations for proportion of Ecological Group 1 benthic infauna are coded BBXX-YYY, where XX is last two digits of year and YYY is station number.

Group	Water property year	Chlorophyll *a*, µg L^{-1}	Dissolved oxygen, mg L^{-1}	Salinity	Temperature	Total N, µg L^{-1}	Total P, µg L^{-1}	Benthic station	Proportion EG1 species
1	2013	5.05	6.81	27.382	22.508	336.63	57.443	BB14-097	0.159
1	2013	5.05	6.81	27.382	22.508	336.63	57.443	BB14-093	0.389
1	2013	5.05	6.81	27.382	22.508	336.63	57.443	BB14-098	0.59
1	2013	5.05	6.81	27.382	22.508	336.63	57.443	BB14-095	0.161
2	2011	8.2	5.89	28	25.1	434.2	65.1	BB12-084	0.1
2	2011	8.2	5.89	28	25.1	434.2	65.1	BB12-083	0.25
2	2012	4.75	6.1	28	25.2	398.2	76.3	BB13-084	0.329
2	2012	4.75	6.1	28	25.2	398.2	76.3	BB13-083	0.518
3	2011	9.95	5.72	28.1	24.6	473.5	62.6	BB12-088	0.171
3	2011	9.95	5.72	28.1	24.6	473.5	62.6	BB12-096	0.698
3	2011	9.95	5.72	28.1	24.6	473.5	62.6	BB12-094	0.13
3	2011	9.95	5.72	28.1	24.6	473.5	62.6	BB12-091	0.122
3	2011	9.95	5.72	28.1	24.6	473.5	62.6	BB12-079	0.191
3	2012	7.82	6.09	29	24.8	394.9	80.3	BB13-096	0.563
3	2012	7.82	6.09	29	24.8	394.9	80.3	BB13-088	0.26
3	2012	7.82	6.09	29	24.8	394.9	80.3	BB13-094	0.291
3	2012	7.82	6.09	29	24.8	394.9	80.3	BB13-079	0.363
3	2012	7.82	6.09	29	24.8	394.9	80.3	BB13-091	0.153
3	2013	5.59	6.57	27.4	22.3	329	60	BB14-096	0.547
3	2013	5.59	6.57	27.4	22.3	329	60	BB14-094	0.162
3	2013	5.59	6.57	27.4	22.3	329	60	BB14-088	0.094
3	2013	5.59	6.57	27.4	22.3	329	60	BB14-079	0.275
3	2013	5.59	6.57	27.4	22.3	329	60	BB14-091	0.314
4	2011	13.9	5.68	27.8	25.319	581.2	106.7	BB12-074	0.155
4	2011	13.9	5.68	27.8	25.319	581.2	106.7	BB12-085	0.244
4	2011	13.9	5.68	27.8	25.319	581.2	106.7	BB12-089	0.114
4	2011	13.9	5.68	27.8	25.319	581.2	106.7	BB12-075	0.185
4	2012	6	5.97	26.6	25.3	465.6	73.9	BB13-074	0.173
4	2012	6	5.97	26.6	25.3	465.6	73.9	BB13-089	0.212
4	2012	6	5.97	26.6	25.3	465.6	73.9	BB13-085	0.244

4	2012	6	5.97	26.6	25.3	465.6	73.9	BB13-075	0.186
5	2013	5.33	6.41	27.122	23.733	494.4	82.547	BB14-073	0.064
5	2013	5.33	6.41	27.122	23.733	494.4	82.547	BB14-062	0.234
5	2013	5.33	6.41	27.122	23.733	494.4	82.547	BB14-069	0.239
5	2013	5.33	6.41	27.122	23.733	494.4	82.547	BB14-068	0.167
6	2011	8.2	5.4	27.7	25.4	595.2	94.4	BB12-060	0.15
6	2012	6.24	6.46	26.6	25.3	454.1	67.3	BB13-060	0.109
6	2012	6.24	6.46	26.6	25.3	454.1	67.3	BB13-059	0.257
7	2011	10.86	5.67	27.4	24.5	570.7	93.3	BB12-042	0.14
7	2011	10.86	5.67	27.4	24.5	570.7	93.3	BB12-063	0.12
7	2011	10.86	5.67	27.4	24.5	570.7	93.3	BB12-059	0.098
7	2011	10.86	5.67	27.4	24.5	570.7	93.3	BB12-039	0.132
7	2011	10.86	5.67	27.4	24.5	570.7	93.3	BB12-037	0.152
7	2011	10.86	5.67	27.4	24.5	570.7	93.3	BB12-056	0.173
7	2011	10.86	5.67	27.4	24.5	570.7	93.3	BB12-052	0.286
7	2011	10.86	5.67	27.4	24.5	570.7	93.3	BB12-051	0.132
7	2012	4.38	5.98	26.8	25.5	439.5	49.7	BB13-056	0.191
7	2012	4.38	5.98	26.8	25.5	439.5	49.7	BB13-052	0.157
7	2012	4.38	5.98	26.8	25.5	439.5	49.7	BB13-051	0.38
7	2012	4.38	5.98	26.8	25.5	439.5	49.7	BB13-042	0.11
7	2012	4.38	5.98	26.8	25.5	439.5	49.7	BB13-063	0.223
7	2012	4.38	5.98	26.8	25.5	439.5	49.7	BB13-039	0.092
7	2012	4.38	5.98	26.8	25.5	439.5	49.7	BB13-037	0.301
7	2013	4.84	5.66	26.8	24.1	488.2	66.6	BB14-056	0.281
7	2013	4.84	5.66	26.8	24.1	488.2	66.6	BB14-051	0.363
7	2013	4.84	5.66	26.8	24.1	488.2	66.6	BB14-039	0.077
7	2013	4.84	5.66	26.8	24.1	488.2	66.6	BB14-037	0.558
7	2013	4.84	5.66	26.8	24.1	488.2	66.6	BB14-052	0.211
7	2013	4.84	5.66	26.8	24.1	488.2	66.6	BB14-063	0.339
7	2013	4.84	5.66	26.8	24.1	488.2	66.6	BB14-042	0.168
8	2011	6.66	6.24	27.8	24.5	413	42.5	BB12-050	0.088
8	2011	6.66	6.24	27.8	24.5	413	42.5	BB12-048	0.254
8	2011	6.66	6.24	27.8	24.5	413	42.5	BB12-045	0.169
8	2011	6.66	6.24	27.8	24.5	413	42.5	BB12-043	0.16
8	2011	6.66	6.24	27.8	24.5	413	42.5	BB12-040	0.511
8	2012	7.93	7.8	29.7	24.4	428.8	53.3	BB13-048	0.222
8	2012	7.93	7.8	29.7	24.4	428.8	53.3	BB13-045	0.231
8	2012	7.93	7.8	29.7	24.4	428.8	53.3	BB13-043	0.165

8	2012	7.93	7.8	29.7	24.4	428.8	53.3	BB13-040	0.522
8	2012	7.93	7.8	29.7	24.4	428.8	53.3	BB13-050	0.343
8	2013	4.79	7.19	26.3	23	342.4	41.8	BB14-048	0.203
8	2013	4.79	7.19	26.3	23	342.4	41.8	BB14-050	0.453
8	2013	4.79	7.19	26.3	23	342.4	41.8	BB14-045	0.159
8	2013	4.79	7.19	26.3	23	342.4	41.8	BB14-043	0.262
8	2013	4.79	7.19	26.3	23	342.4	41.8	BB14-040	0.805
9	2013	6.62	6.66	25.6	23.5	432	51.4	BB14-054	0.462
9	2013	6.62	6.66	25.6	23.5	432	51.4	BB14-053	0.313
9	2013	6.62	6.66	25.6	23.5	432	51.4	BB14-049	0.206
9	2013	6.62	6.66	25.6	23.5	432	51.4	BB14-058	0.396
10	2011	6.45	6.28	24.7	25	460	39.2	BB12-036	0.131
10	2011	6.45	6.28	24.7	25	460	39.2	BB12-034	0.262
10	2011	6.45	6.28	24.7	25	460	39.2	BB12-026	0.185
10	2012	9.41	6.52	25.6	25.2	503.6	52.2	BB13-036	0.266
10	2012	9.41	6.52	25.6	25.2	503.6	52.2	BB13-034	0.19
10	2012	9.41	6.52	25.6	25.2	503.6	52.2	BB13-026	0.23
10	2013	7.48	6.86	23.9	22.7	545.4	61.9	BB14-036	0.406
10	2013	7.48	6.86	23.9	22.7	545.4	61.9	BB14-034	0.393
10	2013	7.48	6.86	23.9	22.7	545.4	61.9	BB14-026	0.172
11	2011	7.92	6.0133	22.7	25.078	492.5	32.2	BB12-029	0.234
11	2011	7.92	6.0133	22.7	25.078	492.5	32.2	BB12-028	0.232
11	2011	7.92	6.0133	22.7	25.078	492.5	32.2	BB12-021	0.166
11	2011	7.92	6.0133	22.7	25.078	492.5	32.2	BB12-033	0.142
11	2011	7.92	6.0133	22.7	25.078	492.5	32.2	BB12-030	0.173
12	2012	12.96	6.69	20	25.4	647.2	58.2	BB13-032	0.165
12	2012	12.96	6.69	20	25.4	647.2	58.2	BB13-031	0.084
12	2012	12.96	6.69	20	25.4	647.2	58.2	BB13-022	0.219
12	2012	12.96	6.69	20	25.4	647.2	58.2	BB13-017	0.195
12	2012	12.96	6.69	20	25.4	647.2	58.2	BB13-016	0.15
13	2011	9.04	6.96	17.7	25.5	699.2	34.5	BB12-020	0.089
14	2012	16.5	6.67	16.7	25.4	692.2	53	BB13-024	0
14	2012	16.5	6.67	16.7	25.4	692.2	53	BB13-015	0
14	2013	13.37	6.3	17.1	23.6	820.2	64.3	BB14-024	0.043
14	2013	13.37	6.3	17.1	23.6	820.2	64.3	BB14-015	0.015
15	2011	11.78	6.59	18.8	25.4	633	30.5	BB12-001	0.099
15	2012	12.08	6.99	18.3	25.4	646.2	50.8	BB13-001	0.043
15	2013	12.11	6.93	18.5	23	829	74.6	BB14-001	0.173

16	2011	10.26	6.89	18.4	25.3	693.8	32	BB12-008	0.163
16	2011	10.26	6.89	18.4	25.3	693.8	32	BB12-004	0.217
16	2011	10.26	6.89	18.4	25.3	693.8	32	BB12-002	0.101
16	2012	11.25	7.22	17.7	25.3	722.8	55.4	BB13-008	0.096
16	2012	11.25	7.22	17.7	25.3	722.8	55.4	BB13-004	0.122
16	2012	11.25	7.22	17.7	25.3	722.8	55.4	BB13-002	0.139
17	2011	11.49	6.516	20.4	25	633.8	41.4	BB12-011	0.086
17	2012	12.04	6.51	20.5	25.4	633	59.4	BB13-011	0.32
17	2013	9.05	6.53	18.1	22.9	630.2	61.3	BB14-011	0.048

Table 3. Initial model, stepwise linear regression of proportion of EG1 species against water properties.

```
Unforced Variables: chlorophyll DO salinity temperature total_N total_P
  P to Enter 0.0500
  P to Exit  0.0500
```

Step	Variable	Coefficient	T	P	R^2	MSE
1	Constant	0.38338	0.76		0.2459	0.01655
	chlorophyll	-4.092E-03	-0.60	0.5486		
	DO	0.03072	1.13	0.2613		
	salinity	6.147E-03	0.68	0.4995		
	temperature	-0.01336	-0.85	0.3987		
	total_N	-3.030E-04	-1.02	0.3087		
	total_P	4.509E-05	0.04	0.9697		
2	Constant	0.38461	0.76		0.2459	0.01639
	chlorophyll	-4.020E-03	-0.62	0.5378		
	DO	0.03035	1.20	0.2320		
	salinity	6.396E-03	1.03	0.3075		
	temperature	-0.01361	-0.95	0.3418		
	total_N	-2.967E-04	-1.21	0.2289		
3	Constant	0.46183	0.95		0.2433	0.01630
	DO	0.02822	1.13	0.2604		
	salinity	5.620E-03	0.92	0.3583		
	temperature	-0.01476	-1.05	0.2969		
	total_N	-3.934E-04	-2.09	0.0390		
4	Constant	0.72626	1.85		0.2374	0.01628
	DO	0.01796	0.80	0.4226		
	temperature	-0.01424	-1.01	0.3136		
	total_N	-5.329E-04	-4.75	0.0000		
5	Constant	0.92244	2.99		0.2330	0.01623
	temperature	-0.01784	-1.34	0.1830		
	total_N	-5.207E-04	-4.69	0.0000		
6	Constant	0.51551	9.79		0.2207	0.01634
	total_N	-5.794E-04	-5.66	0.0000		

Resulting Stepwise Model

Variable	Coefficient	Std Error	T	P	VIF
Constant	0.51551	0.05265	9.79	0.0000	
total_N	-5.794E-04	1.024E-04	-5.66	0.0000	1.0

Cases Included	115	R^2	0.2207	MSE	0.01634
Missing Cases	0	Adjusted R^2	0.2138	SD	0.12784

Variables Not in the Model

Variable	Correlations Multiple	Partial	T	P
chlorophyll	0.7801	-0.0515	-0.55	0.5863
DO	0.0031	0.1116	1.19	0.2372
salinity	0.7892	0.0199	0.21	0.8334
temperature	0.3945	-0.1256	-1.34	0.1830

```
total_P        0.1071    0.0125    0.13   0.8947
```

Table 4. Average values of parameters in each station group used for final model development.

Group	Chlorophyll *a*, µg L^{-1}	Dissolved oxygen, mg L^{-1}	Salinity	Temperature	Total N, µg L^{-1}	Total P, µg L^{-1}	Proportion EG1 species
1	5.05	6.81	27.4	22.5	336.6	57.4	0.325
2	6.48	6.00	28.0	25.1	416.3	70.7	0.299
3	7.79	6.13	28.2	23.9	399.1	67.7	0.289
4	9.94	5.83	27.2	25.3	523.4	90.3	0.186
5	5.33	6.41	27.1	23.7	494.4	82.5	0.176
6	7.22	5.93	27.2	25.4	524.6	80.8	0.172
7	6.69	5.77	27.0	24.7	499.5	69.9	0.213
8	6.46	7.07	27.9	23.9	394.7	45.8	0.303
9	6.62	6.67	25.6	23.5	432.1	51.3	0.344
10	7.78	6.55	24.8	24.3	503.0	51.1	0.248
11	7.92	6.01	22.7	25.1	492.5	32.2	0.189
12	12.95	6.69	20.0	25.4	647.2	58.2	0.163
13	9.04	6.96	17.7	25.5	699.3	34.5	0.089
14	14.94	6.49	16.9	24.5	756.1	58.6	0.014
15	11.99	6.84	18.5	24.6	702.7	52.0	0.105
16	10.75	7.05	18.0	25.4	708.2	43.7	0.14
17	10.86	6.52	19.6	24.5	632.4	54.0	0.151

Table 5. Final model, stepwise linear regression of proportion of sensitive species

Unforced Variables: Chlorophyll Salinity Total_N
P to Enter 0.0500
P to Exit 0.0500

Step	Variable	Coefficient	T	P	R^2	MSE
1	Constant	0.85648	3.90		0.8748	1.273E-03
	Chlorophyll	1.523E-03	0.24	0.8110		
	Salinity	-7.473E-03	-1.39	0.1887		
	Total_N	-9.126E-04	-4.53	0.0006		
2	Constant	0.86170	4.08		0.8742	1.187E-03
	Salinity	-7.655E-03	-1.49	0.1595		
	Total_N	-8.897E-04	-5.16	0.0001		
3	Constant	0.55283	14.34		0.8544	1.283E-03
	Total_N	-6.540E-04	-9.38	0.0000		

Resulting Stepwise Model

Variable	Coefficient	Std Error	T	P	VIF
Constant	0.55283	0.03856	14.34	0.0000	
Total_N	-6.540E-04	6.971E-05	-9.38	0.0000	1.0

Cases Included	17	R^2	0.8544	MSE	1.283E-03
Missing Cases	0	Adjusted R^2	0.8447	SD	0.03581

Variables Not in the Model

	Correlations			
Variable	Multiple	Partial	T	P
Chlorophyll	0.8574	0.1133	0.43	0.6762
Salinity	0.9211	-0.3690	-1.49	0.1595

Table 6. Station locations and water properties for 2016. All samples collected on 07/07/2016.

					Surface					Bottom				
Station	Latitude	Longitude	Time	Depth	Salinity	DO (mg L^{-1})	DO (%)	pH	Temp	Salinity	DO (mg L^{-1})	DO (%)	pH	Temp
BB16-001	40.03931	-74.054	13:05	1.5	24.13	6.63	94.9	7.77	26.8	24.67	5.86	82.4	7.66	25.6
BB16-002	39.94816	-74.1016	12:31	1.8	22.34	6.66	96.5	8.04	28.1	22.32	6.87	99.4	8.04	28
BB16-003	39.93744	-74.11	12:14	1.8	22.82	6.48	92.7	7.92	27.2	22.81	6.13	87.7	7.92	27.2
BB16-004	39.93247	-74.1407	11:50	2	19.36	7.11	100.8	8.03	27.8	23.34	5.56	78.4	7.78	26.2
BB16-005	39.85253	-74.102	11:17	1.5	27.37	6.51	95.2	7.98	27	27.9	6.22	90.5	8	26.5
BB16-006	39.8011	-74.1573	10:57	2.8	30.05	6.67	96.6	7.86	25.5	31.38	6.25	82.6	7.71	19.9
BB16-007	39.74254	-74.1462	10:32	1.4	30.7	5.51	81.7	7.87	26.7	30.7	5.19	77.0	7.86	26.7
BB16-008	39.71856	-74.1717	10:14	1.1	29.68	4.86	73.1	7.63	27.9	29.74	4.36	65.4	7.64	27.7
BB16-009	39.66078	-74.2068	9:47	2.1	30	5.86	87.3	7.98	27.2	30.02	5.71	85.1	7.96	27.2
BB16-010	39.64789	-74.2195	9:23	0.6	24.05	4.86	70.4	7.49	27.5	28.38	4.51	66.5	7.59	27.1
BB16-011	39.5953	-74.2515	8:56	1.7	30.83	5.25	77.3	7.73	26.2	30.83	5.02	73.4	7.68	25.8
BB16-012	39.58127	-74.2681	8:40	1.4	30.84	5.76	82.8	7.7	24.8	30.84	5.77	82.9	7.7	24.8
BB16-013	39.56885	-74.3247	7:40	1.2	30.08	6.41	93.5	7.77	25.9	30.6	5.56	79.7	7.66	24.7
BB16-014	39.51102	-74.2973	8:05	6.7	31.15	5.79	79.2	7.6	21.9	31.25	5.51	74.5	7.54	21.2

Table 7. Sediment properties for 2016.

Station	% gravel	% sand	% silt	%C	%N	%P
BB16-001	0.13	93.18	6.69	3.805E-01	4.150E-02	7.49E-03
BB16-002	0	60.22	39.78	1.367E+00	1.213E-01	4.45E-02
BB16-003	0	33.77	65.67	2.896E+00	2.272E-01	6.39E-02
BB16-004	0	4.44	95.56	6.161E+00	4.885E-01	1.05E-01
BB16-005	0	95.02	4.98	2.048E-01	1.698E-02	8.57E-03
BB16-006	0.12	65.4	34.48	8.266E-01	9.700E-02	6.17E-02
BB16-007	0	94.17	5.82	1.711E-01	1.472E-02	1.19E-02
BB16-008	0	6.38	93.62	3.034E+00	2.660E-01	6.41E-02
BB16-009	0	78.94	21.07	5.936E-01	7.121E-02	2.92E-02
BB16-010	0	9.28	90.72	4.344E+00	3.341E-01	8.94E-02
BB16-011	0.03	65.29	34.67	8.433E-01	8.141E-02	5.75E-02
BB16-012	0	81.99	18.02	5.413E-01	5.916E-02	4.19E-02
BB16-013	0.08	45.18	54.74	1.019E+00	5.303E-02	7.90E-02
BB16-014	0.07	98.44	1.5	1.713E-02	4.089E-04	1.21E-03

Table 8. Species assignments to Ecological Groups and their presence or absence in samples collected in different years.

		Present in:			
Species	Ecological Group	2012	2013	2014	2016
Acanthohaustorius millsi	1	yes	yes	yes	
Acteocina canaliculata	1	yes	yes	yes	yes
Aligena elevata	1	yes	yes	yes	
Amastigos caperatus	1	yes	yes	yes	yes
Ameroculodes spp. complex	1	yes	yes	yes	yes
Ampelisca vadorum	1	yes	yes	yes	yes
Ampelisca verrilli	1	yes	yes	yes	yes
Ampharete oculata	1	yes	yes	yes	yes
Amphitrite ornata	1	yes	yes	yes	yes
Arabella iricolor	1	yes	yes	yes	yes
Aricidea (Aricidea) wassi	1	yes	yes	yes	
Batea catharinensis	1	yes	yes	yes	yes
Bittiolum alternatum	1	yes	yes	yes	yes
Callipallene brevirostris	1	yes	yes	yes	yes
Carinomella lactea	1	yes	yes	yes	yes
Ceriantheopsis americanus	1	yes	yes	yes	yes
Clymenella torquata	1	yes	yes	yes	yes
Clymenella zonalis	1	yes	yes	yes	yes
Costoanachis avara	1	yes	yes	yes	
Cymadusa compta	1	yes	yes	yes	yes
Diopatra cuprea	1	yes	yes	yes	yes
Elasmopus levis	1	yes	yes	yes	yes
Ensis directus	1	yes	yes	yes	yes
Eobrolgus spinosus	1	yes	yes	yes	yes
Epitonium rupicola	1	yes	yes	yes	
Gammarus mucronatus	1	yes	yes	yes	yes
Globosolembos smithi	1	yes			
Havelockia scabra	1	yes	yes	yes	
Japonactaeon punctostriatus	1	yes	yes	yes	yes
Leptosynapta tenuis	1	yes	yes	yes	yes
Lyonsia hyalina	1	yes	yes	yes	yes
Melita nitida	1	yes	yes	yes	
Molgula manhattensis	1			yes	
Nucula proxima	1	yes	yes	yes	yes
Oxydromus obscurus	1	yes	yes		
Oxyurostylis smithi	1	yes	yes	yes	yes
Parahaustorius longimerus	1	yes	yes	yes	yes
Parasabella microphthalma	1	yes	yes	yes	yes
Pectinaria gouldii	1	yes	yes	yes	yes
Pentamera pulcherrima	1	yes	yes	yes	yes
Pista cristata	1	yes	yes	yes	yes
Pista palmata	1	yes	yes	yes	yes
Polygordius jouinae	1	yes	yes	yes	yes

Protohaustorius cf. deichmannae	1	yes	yes	yes	
Ptilanthura tenuis	1	yes	yes	yes	yes
Rhepoxynius hudsoni	1	yes	yes	yes	
Rudilemboides naglei	1	yes	yes	yes	yes
Sabaco elongatus	1	yes	yes	yes	yes
Saccoglossus kowalevskii	1	yes	yes	yes	yes
Solemya velum	1	yes	yes	yes	yes
Spisula solidissima	1	yes	yes	yes	yes
Turbonilla interrupta	1	yes	yes	yes	yes
Unciola dissimilis	1	yes	yes	yes	yes
Unciola irrorata	1	yes	yes	yes	yes
Unciola serrata	1	yes	yes	yes	yes
Upogebia affinis	1	yes			
Yoldia limatula	1	yes	yes	yes	yes
Americhelidium americanum	2	yes		yes	yes
Ameritella agilis	2	yes	yes	yes	yes
Ampithoe longimana	2	yes	yes	yes	yes
Anoplodactylus petiolatus	2	yes	yes	yes	yes
Aricidea (Acmira) catherinae	2	yes	yes	yes	yes
Astyris lunata	2		yes	yes	
Brania wellfleetensis	2	yes	yes	yes	yes
Cerebratulus lacteus	2	yes	yes	yes	yes
Chiridotea coeca	2	yes	yes	yes	yes
Cyathura burbancki	2	yes	yes	yes	yes
Cyclaspis varians	2	yes	yes	yes	yes
Drilonereis longa	2	yes	yes	yes	yes
Edotia triloba	2	yes	yes	yes	yes
Edwardsia elegans	2	yes	yes	yes	yes
Erinaceusyllis erinaceus	2	yes	yes	yes	yes
Eulimastoma engonium	2	yes	yes	yes	yes
Eumida sanguinea	2	yes	yes	yes	yes
Euplana gracilis	2	yes	yes	yes	yes
Exogone (Exogone) dispar	2	yes	yes	yes	yes
Gemma gemma	2	yes	yes	yes	yes
Glycera americana	2	yes	yes	yes	yes
Glycera dibranchiata	2	yes	yes	yes	yes
Glycinde multidens	2	yes	yes	yes	yes
Haloclava producta	2	yes	yes	yes	yes
Haminoea solitaria	2	yes	yes	yes	yes
Hargeria rapax	2	yes	yes	yes	
Harmothoe extenuata	2	yes	yes	yes	yes
Idotea balthica	2	yes	yes	yes	yes
Idunella barnardi	2	yes	yes	yes	yes
Leucon americanus	2	yes	yes	yes	yes
Lysianopsis alba	2	yes	yes	yes	yes
Lysilla alba	2	yes	yes	yes	
Marenzelleria viridis	2	yes	yes	yes	
Marphysa bellii	2	yes	yes	yes	

Melinna maculata	2	yes	yes	yes	yes
Mercenaria mercenaria	2	yes	yes	yes	yes
Microphthalmus aggregatus	2	yes	yes	yes	
Microphthalmus sczelkowii	2		yes	yes	yes
Mya arenaria	2	yes	yes	yes	
Nephtys incisa	2	yes	yes	yes	yes
Nephtys picta	2	yes	yes	yes	
Notocirrus spinifera	2	yes	yes	yes	yes
Owenia fusiformis	2	yes	yes	yes	
Pagurus longicarpus	2	yes	yes	yes	
Parahesione luteola	2	yes		yes	
Parametopella cypris	2	yes		yes	
Paranaitis speciosa	2	yes	yes	yes	yes
Phascolion (Phascolion) strombus	2	yes	yes	yes	yes
Pholoe minuta	2			yes	yes
Phoronis psammophila	2	yes	yes	yes	
Phrontis vibex	2	yes	yes	yes	yes
Phyllodoce arenae	2	yes	yes	yes	yes
Pionosyllis longocirrata	2		yes	yes	
Podarkeopsis levifuscina	2	yes	yes	yes	yes
Prionospio pygmaeus	2	yes	yes	yes	yes
Proceraea cornuta	2	yes	yes	yes	yes
Salvatoria clavata	2	yes	yes	yes	yes
Scolelepis (Parascolelepis) texana	2	yes	yes	yes	yes
Scolelepis bousfieldi	2	yes	yes	yes	yes
Scoletoma tenuis	2	yes	yes	yes	yes
Sphaerosyllis brevidentata	2	yes	yes	yes	yes
Spiochaetopterus costarum oculatus	2	yes	yes	yes	yes
Stenothoe minuta	2	yes	yes	yes	
Sthenelais boa	2	yes	yes	yes	yes
Stylochus ellipticus	2	yes	yes	yes	
Syllides verrilli	2	yes	yes	yes	
Syllis alternata	2	yes	yes	yes	yes
Tagelus divisus	2	yes	yes	yes	yes
Tanaissus psammophilus	2	yes	yes	yes	yes
Tritia trivittata	2	yes		yes	
Turbellaria sp. A (LTBA)	2	yes		yes	yes
Zygonemertes virescens	2	yes	yes	yes	yes
Alitta succinea	3	yes	yes	yes	yes
Ampelisca abdita	3	yes	yes	yes	yes
Amphiporus bioculatus	3	yes	yes	yes	yes
Amphiporus ochraceus	3	yes	yes	yes	yes
Apocorophium acutum	3	yes	yes	yes	
Arenicola cristata	3	yes	yes	yes	
Carinoma tremaphoros	3	yes	yes	yes	yes
Caulleriella venefica	3	yes	yes	yes	yes
Cephalothrix spiralis	3	yes	yes	yes	yes
Cirrophorus sp. B	3	yes	yes	yes	

Erichsonella attenuata	3	yes	yes	yes	
Erichsonella filiformis	3	yes	yes	yes	yes
Hypereteone foliosa	3	yes	yes	yes	
Leptocheirus plumulosus	3		yes	yes	
Lineus ruber	3	yes	yes	yes	
Loimia medusa	3	yes	yes	yes	
Macoma tenta	3	yes	yes	yes	yes
Mediomastus ambiseta	3	yes	yes	yes	yes
Microdeutopus gryllotalpa	3	yes	yes	yes	yes
Monocorophium acherusicum	3	yes	yes	yes	yes
Monocorophium tuberculatum	3	yes	yes	yes	yes
Mysella planulata	3	yes	yes	yes	
Neanthes arenaceodentata	3	yes	yes	yes	yes
Nemertea sp. 2 (MWRA)	3		yes	yes	yes
Notomastus sp. A Ewing	3	yes	yes	yes	yes
Paraonis fulgens	3	yes	yes	yes	
Platynereis dumerilii	3	yes	yes	yes	yes
Prionospio heterobranchia	3	yes	yes	yes	yes
Pygospio elegans	3	yes	yes	yes	yes
Spio setosa	3	yes	yes	yes	yes
Spiophanes bombyx	3	yes	yes	yes	yes
Streblospio benedicti	3	yes	yes	yes	yes
Tetrastemma elegans	3	yes	yes	yes	
Tetrastemma sp. A	3	yes	yes	yes	
Tritia obsoleta	3	yes	yes	yes	yes
Cossura sp. A Maciolek	4	yes	yes	yes	yes
Dipolydora socialis	4	yes	yes	yes	
Heteromastus filiformis	4	yes	yes	yes	yes
Hypereteone heteropoda	4	yes	yes	yes	yes
Leitoscoloplos robustus	4	yes	yes	yes	yes
Monticellina cf. dorsobranchialis	4	yes	yes	yes	yes
Mulinia lateralis	4	yes	yes	yes	yes
Paraprionospio alata	4	yes	yes	yes	yes
Parougia caeca	4	yes		yes	yes
Polycirrus eximius	4	yes	yes	yes	yes
Polydora cornuta	4	yes	yes	yes	yes
Tharyx sp. A (MWRA)	4	yes	yes	yes	yes
Capitella sp.	5	yes	yes	yes	yes
Oligochaeta	5	yes	yes	yes	yes
Tubificoides sp.	5	yes	yes	yes	yes
Bostrichobranchus pilularis	Not assigned			yes	
Dyspanopeus sayi	Not assigned	yes	yes	yes	
Ianiropsis nr. *tridens*	Not assigned	yes	yes	yes	yes
Pseudohaustorius caroliniensis	Not assigned	yes	yes	yes	
Tricladida sp. B	Not assigned	yes	yes	yes	yes

Table 9.

Benthic station	DEP water quality station	Average total N, µg L^{-1} (range)	Proportion of EG1 taxa collected	Proportion of EG1 taxa predicted
16-001	BB01	768.5 (691.7–826.3)	0.032	0.050
16-002	BB03	694.0 (609.1–778.8)	0.024	0.099
16–003	1632B	724.2 (642.2– 842.8)	0.043	0.079
16–004	BB04a	748.4 (677.8–856.6)	0.013	0.064
16–005	BB06	624.7 (490.2–757.0)	0.382	0.144
16–006	BB07a	476.4 (390.6–562.6)	0.299	0.241
16–007	BB09	539.2 (441.4–702.1)	0.329	0.200
16–008	1675	624.8 (472.7–723.6)	0.107	0.144
16–009	BB10	469.1 (371.1–543.4)	0.188	0.246
16–010	1706	687.3 (579.0–833.4)	0.009	0.104
16–011	1800B	448.7 (317.4–536.0)	0.091	0.260
16–012	BB12	405.3 (342.9–486.0)	0.156	0.288
16–013	1818D	488.1 (392.8–675.9)	0.363	0.234
16–014	BB14	391.6 (314.2–465.7)	0.209	0.297

Table 10. Statistical analyses of residuals from Figure 10B.

Wilcoxon Rank Sum Test for Data in model VS Data for 2016

Variable	Rank Sum	N	U Stat	Mean Rank
Data in model	7539.5	115	869.5	65.6
Data for 2016	845.5	14	740.5	60.4

Normal approximation, Z	0.48
Normal approx, two-tailed P	0.6279

Descriptive Statistics

	Data in model	Data for_2016
N	115	14
Lo 95% CI	-0.0238	-0.0802
Mean	-2.348E-04	-0.0146
Up 95% CI	0.0233	0.0509
SD	0.1276	0.1136
Minimum	-0.2440	-0.1690
Median	-0.0240	-0.0440
Maximum	0.4760	0.2380

Figure 1. Locations of the 17 water quality–benthic infauna station groups listed in Table 1. Each circle has a 2 km radius centered on a water quality station and encompasses nearby benthic infauna stations used for model development.

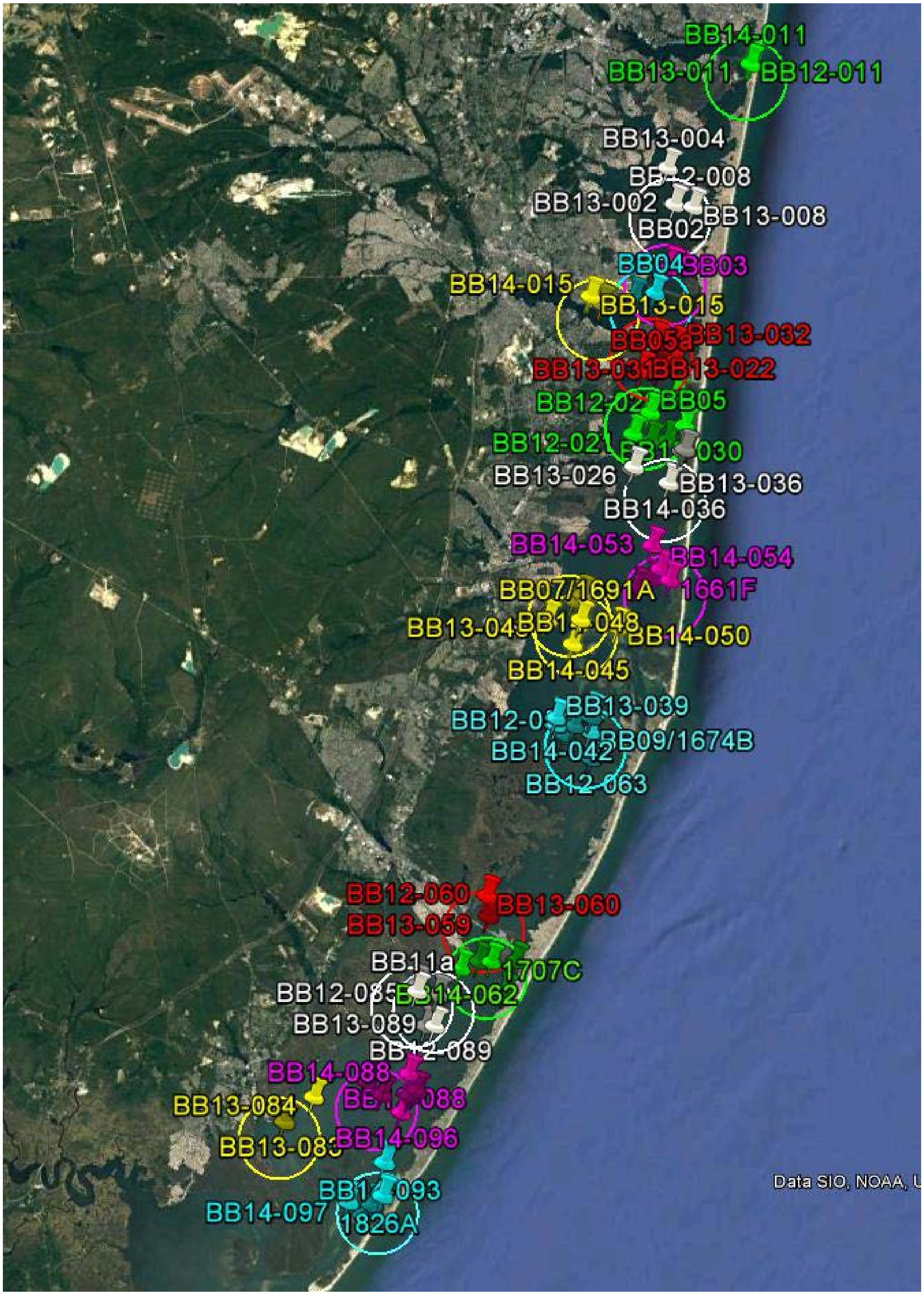
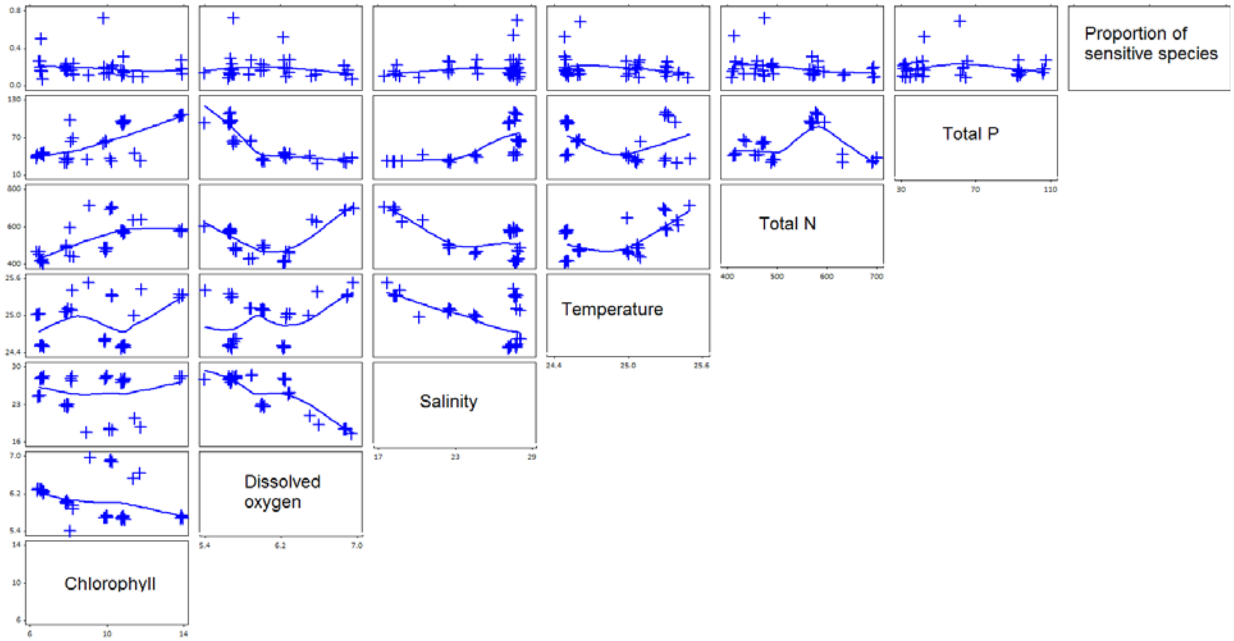


Figure 2. Scatterplot matrices of water properties versus the proportion of EG1 species (top row). Lines are LOESS fits to the data.

(A) 2011 water properties, 2012 benthic invertebrates. N=39



(B) 2012 water properties, 2013 benthic invertebrates. N=40

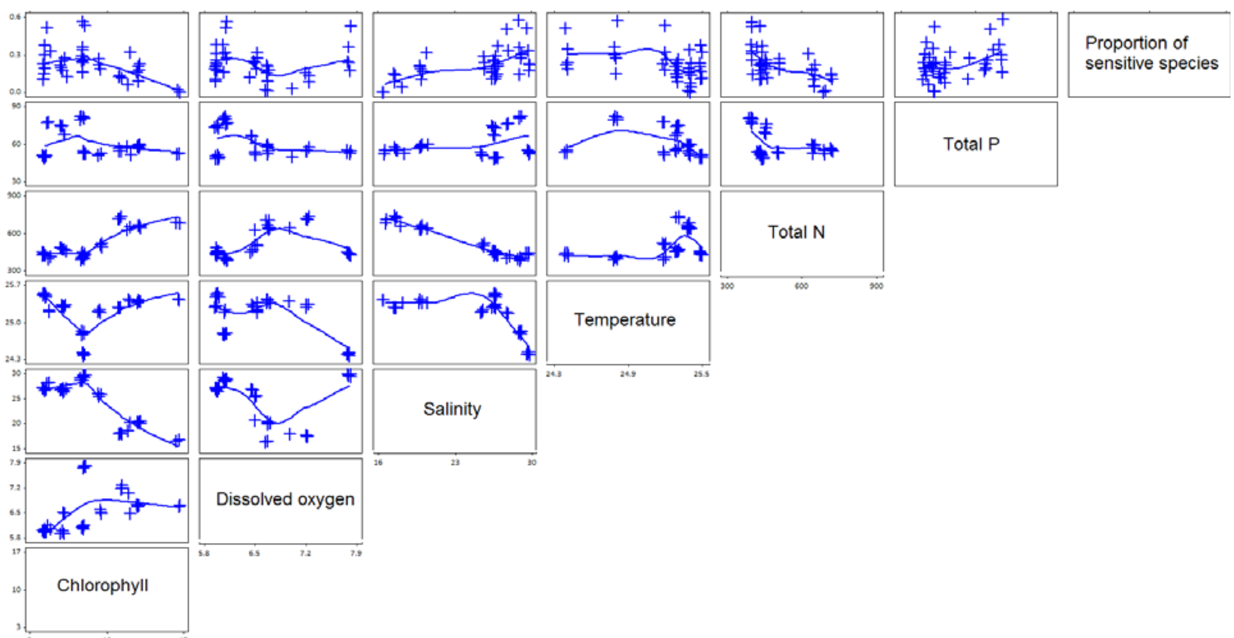
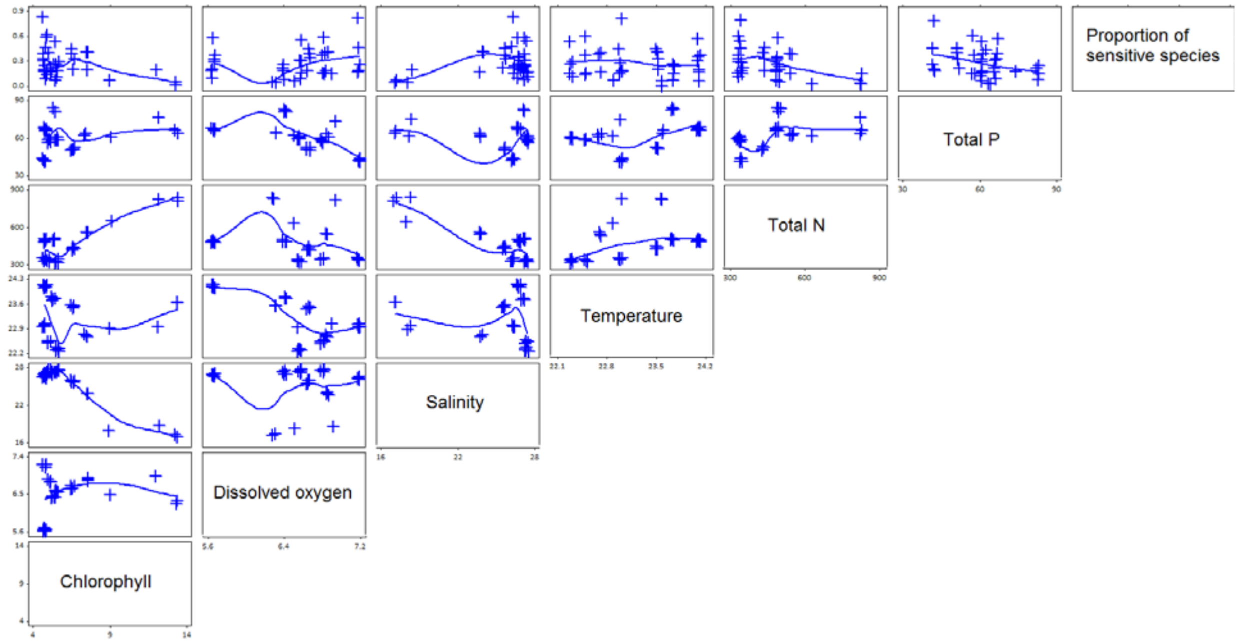


Figure 2, continued.

(C) 2013 water properties, 2014 benthic invertebrates. N=36



(D) All years combined, N=115

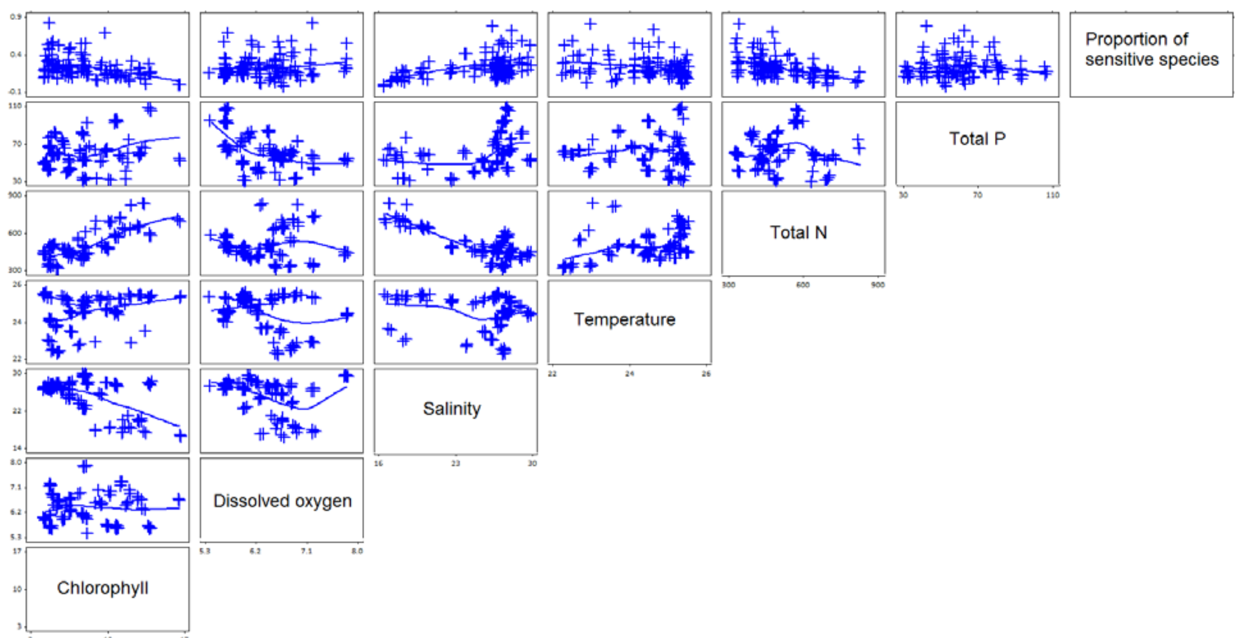
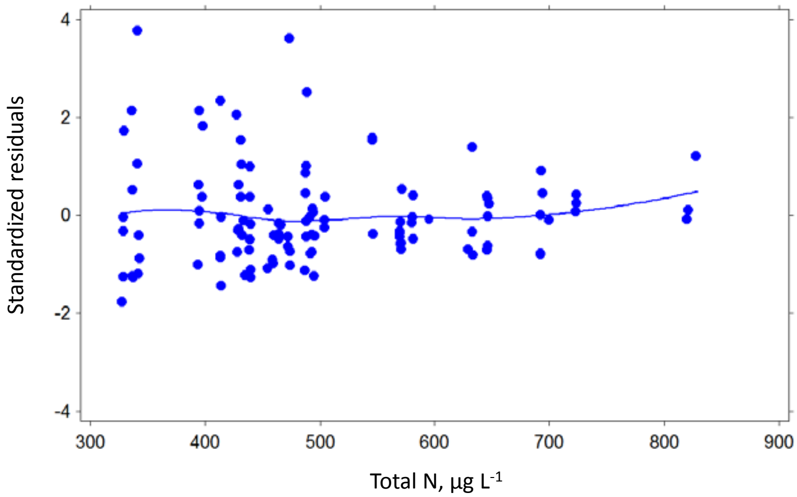
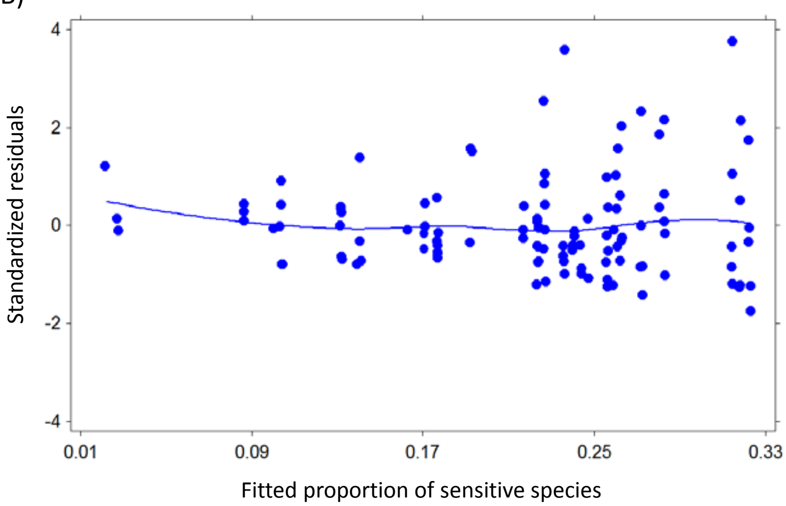


Figure 3. Analysis of residuals from the initial model. Lines in (A) and (B) are LOESS fits.

(A)



(B)



(C)

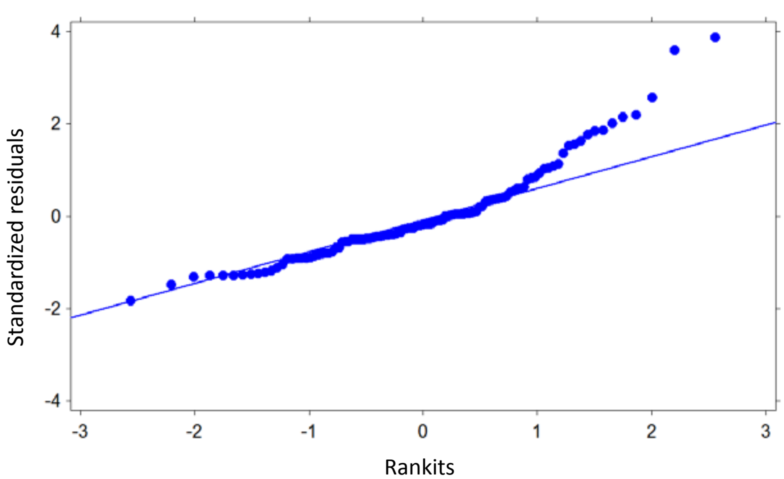


Figure 4. Scatterplot of the relative abundance of EG1 species and total N concentration for all data in the initial model. Data from Table 2, numbers plotted are the Group numbers, multiple entries of the same Group number represent multiple benthic stations near a water sampling station, multiple years, or both. For example, Group 1 has four entries, one for each of the four benthic stations that were near water quality station 1826A (Table 1).

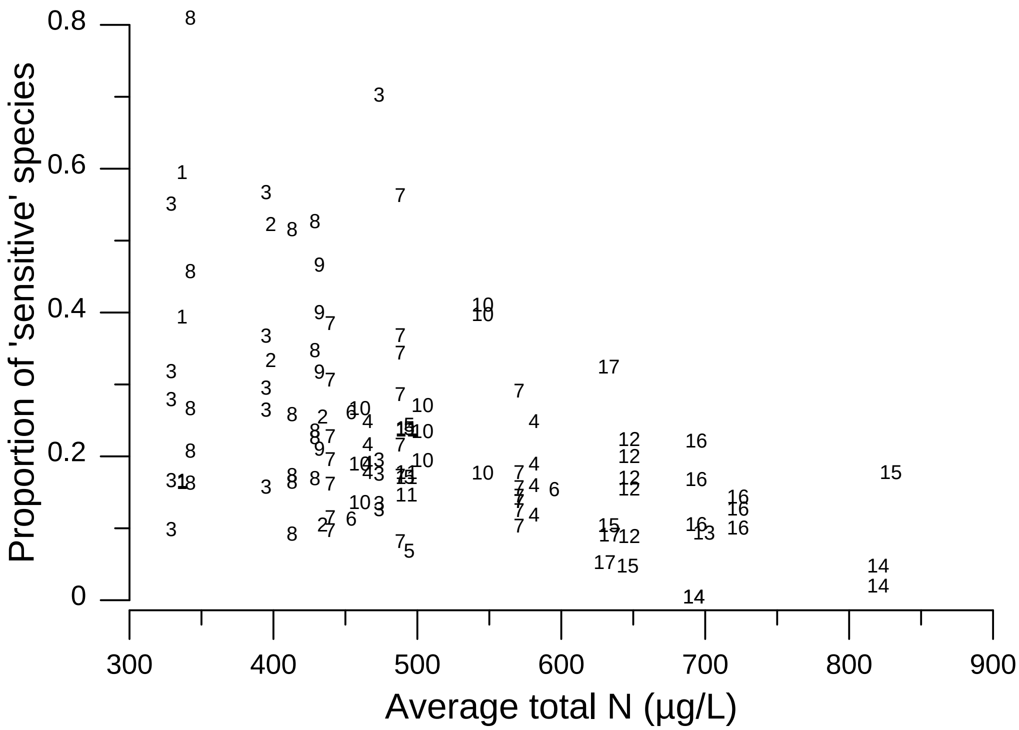


Figure 5. Scatterplot matrices of average values for each of the 17 groups of water properties versus the proportion of EG1 species (top row). Lines are LOESS fits to the data.

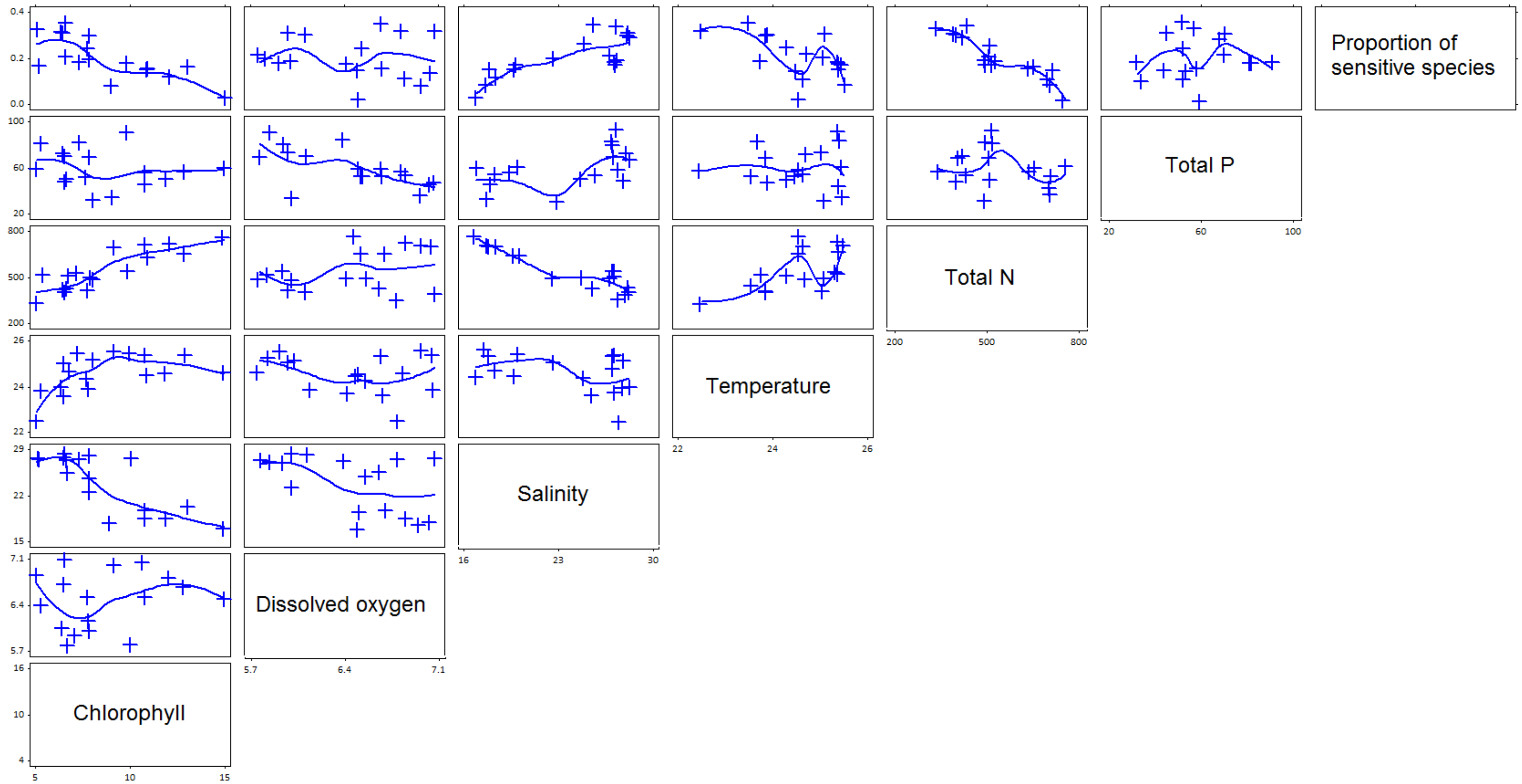
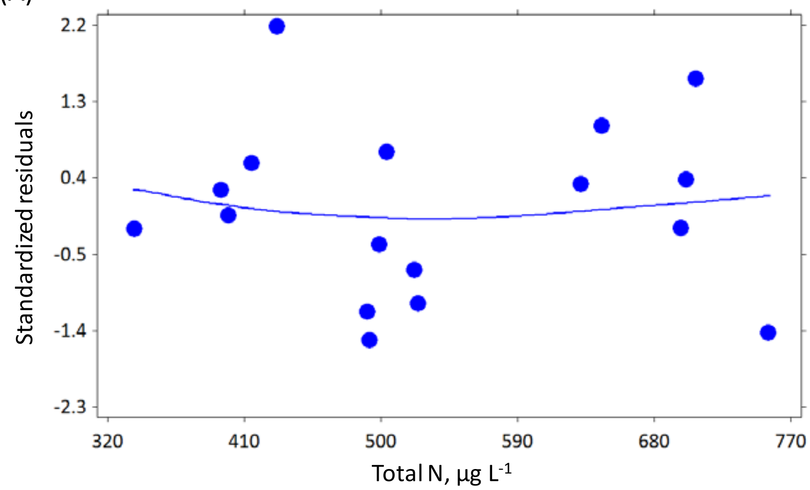
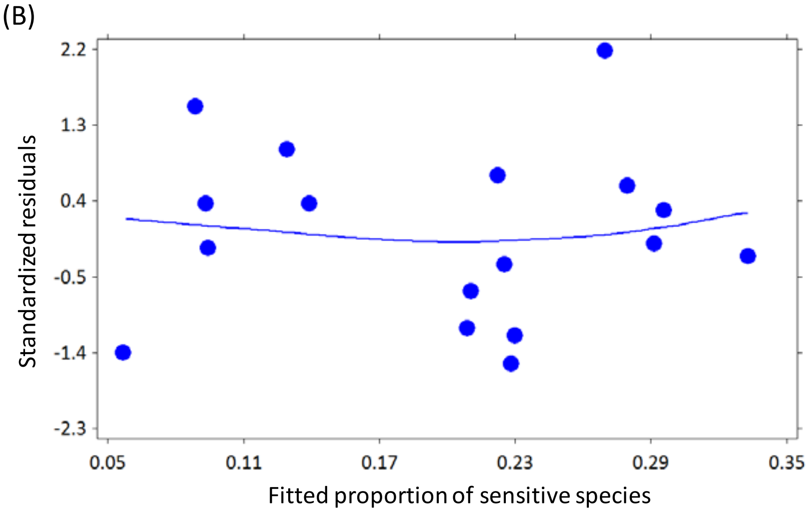


Figure 6. Analysis of residuals for final model. Lines in (A) and (B) are LOESS fits.

(A)



(B)



(C)

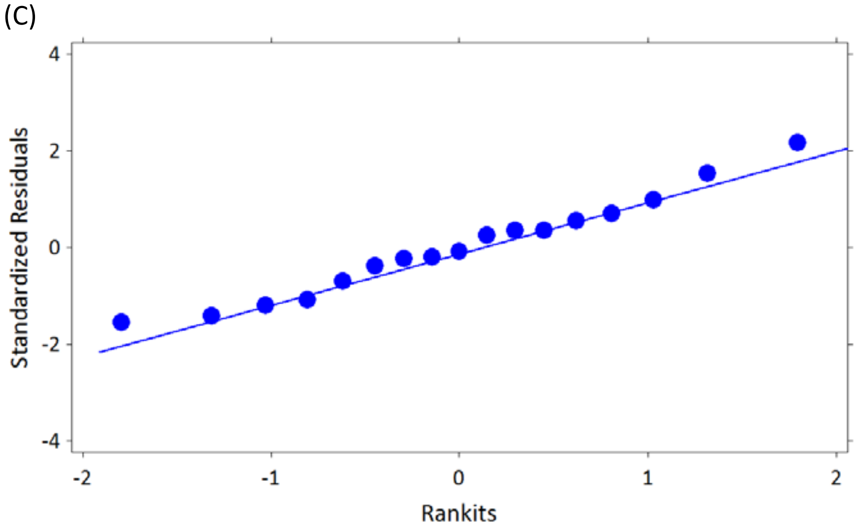


Figure 7. Surface and bottom water properties for benthic stations from 2012 to 2014 used in the model (Table 1) and at the stations sampled in 2016. Open circles are data, Box plots encompass the 75th, 50th (median), and 25th percentiles of the data.

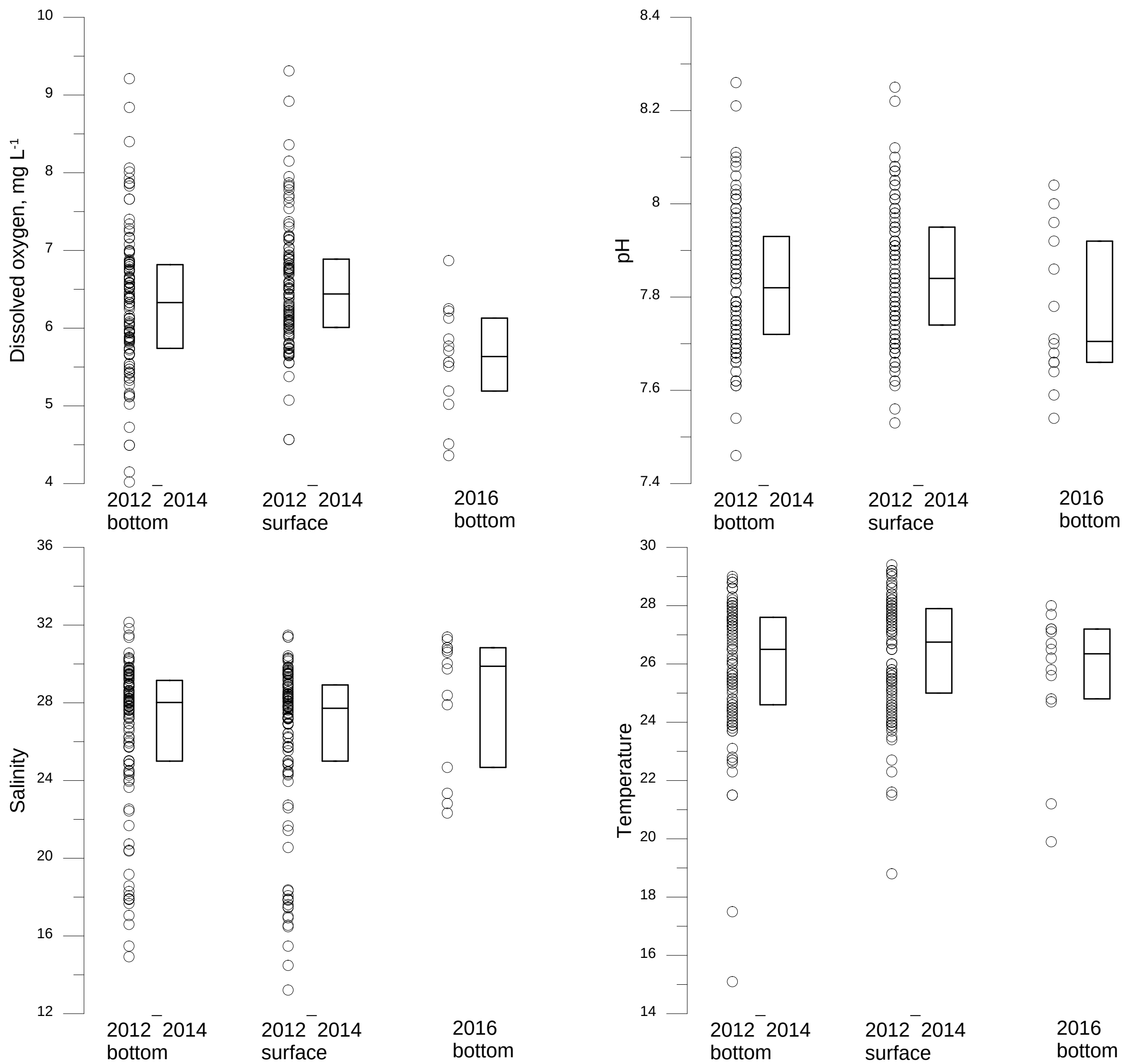


Figure 8. Sediment properties for benthic stations from 2012 to 2014 used in the model (Table 1) and at the stations sampled in 2016. Open circles are data, Box plots encompass the 75th, 50th (median), and 25th percentiles of the data.

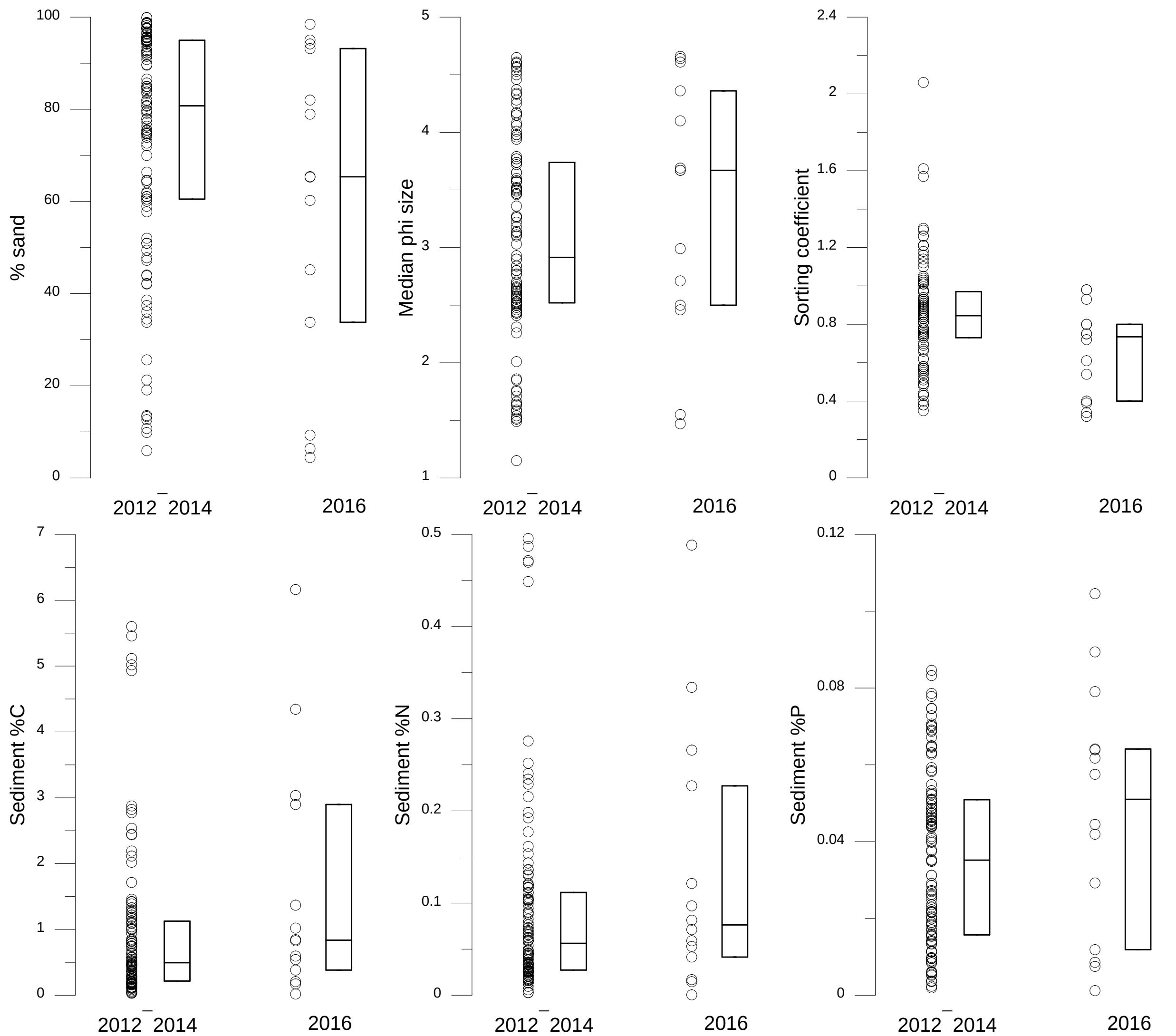


Figure 9. Rarefaction curves of expected total number of species collected versus number of stations sampled in 2016 and in prior years.

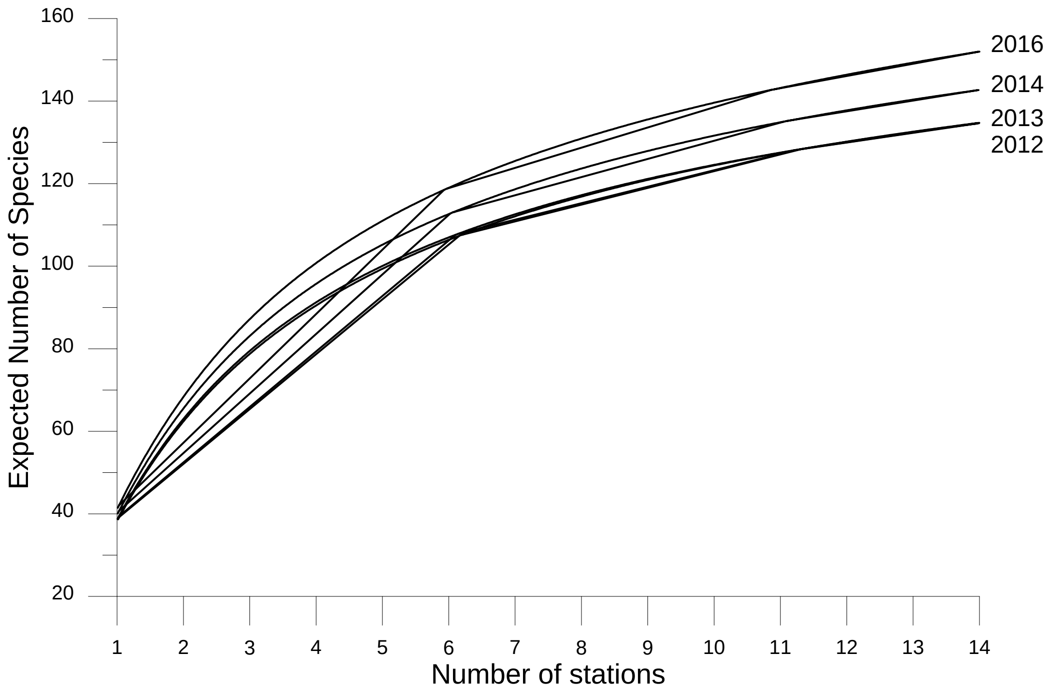
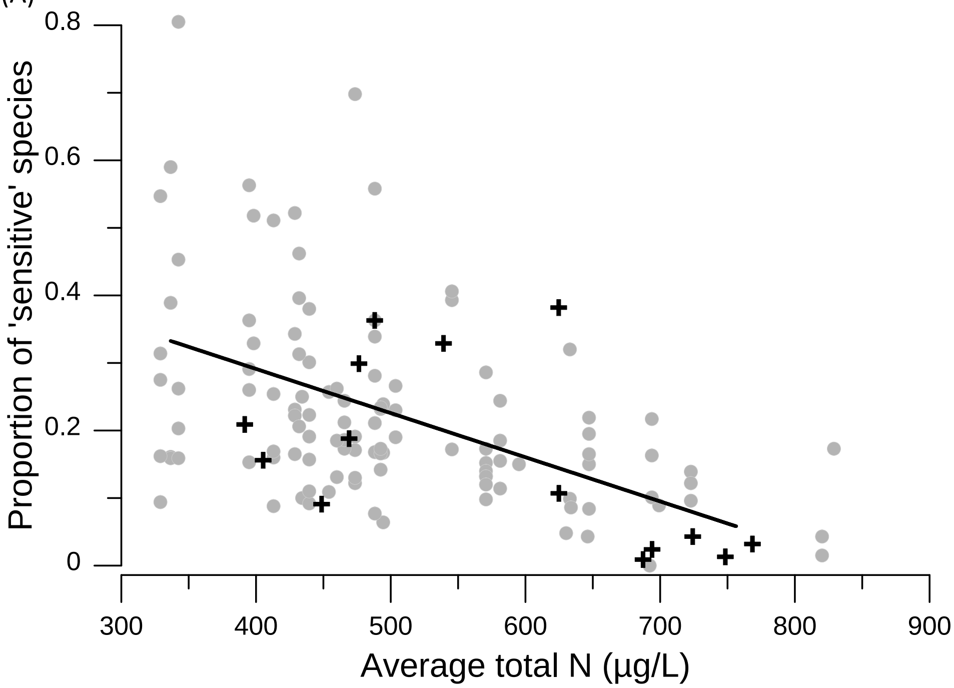


Figure 10. (A) Scatterplot of the relative abundance of EG1 species and total N concentration. Grey filled circles are data used in the linear model (re-plotted data from Figure 4), solid line is linear regression from Table 5. Crosses are data from 2015 for water total N and from 2016 for EG1 species. (B) Open circles are residuals from model prediction for the two data sets, Box plots encompass the 75th, 50th (median), and 25th percentiles of the data.

(A)



(B)