Statistical Analysis of the San Marcos and Comal Springs Aquatic Ecosystems Biomonitoring Datasets

Prepared for:

Edwards Aquifer Authority
900 E. Quincy
San Antonio, Texas 78215

Prepared by:

Beaver Creek Hydrology, LLC.
907 National Avenue
Lexington, Kentucky 40502

TABLE OF CONTENTS

1 INTRODUCTION

In association with increased water utilization from the Edwards Aquifer in south Texas, concern exists for the federally endangered aquatic species residing within the aquifer and at Comal and San Marcos springs. Significant effort to provide adequate spring flow to support these sensitive fish, macroinvertebrate, plant, and salamander species has been initiated via the Edwards Aquifer Habitat Conservation Plan (EARIP 2012). However, the impact of stressors on the communities due to flow, temperature, water chemistry, and recreational pressures still exist. The purpose of this statistical analysis is to perform an exploratory effort to determine the extent of water chemistry variation; what, if any, stressors exist on the populations of concern; and to determine the extent of the variation of species over time and between sampling reaches of Comal and San Marcos spring systems.

In the following report, Beaver Creek Hydrology present the methods, results, discussion, and conclusions related to statistical analyses of biological monitoring data provided by the Edwards Aquifer Authority (EAA). Beyond providing the answers to the questions above, we believe this assessment raises additional questions related to the biology and hydrology in support of compliance with the Long Term Biological Goals (LTBGs) of the Edwards Aquifer Habitat Conservation Plan (EAHCP).

2 METHODS

The primary source of data used in the statistical analyses was collected as part of annual biological monitoring at the headwaters of the Comal and San Marcos systems. The biological monitoring program has been conducted by the same contractors on behalf of the Edwards Aquifer Authority since 2000. Briefly, data are collected at least twice annually on species populations and habitat quality/quantity. Critical period monitoring is triggered after disruptive floods or when low flow thresholds are breached. Standard operating procedures for the biological monitoring program detailing methods and sampling locations can be found with Bio-West (2014a and 2014b). In the present work, data collected from 2000-2015 were used for analyses.

This statistical analysis centers on exploratory multivariate methods. Multivariate statistics are ideal for biological community data as the technique has fewer assumptions regarding data normality (McArdle 2001). As it is rare that a biological collection effort can capture enough data to represent a normal distribution, it is necessary to account for this artifact. Multivariate statistics are also sensitive to trends and associations between various species responses and environmental stressor fluctuations (Ter Braak 1994), thus allowing for an understanding of species/environment trends prior to the development of situational hypotheses (one of the caveats of this project). Especially important is the ability of ordination to parse out natural variation from stressor effects. These techniques can also identify and adapt to issues with co-varying environmental parameters especially when these factors are often interactive and their effects are multiplicative (McCune 2006, 2009).

Ordination's most effective principle lies with its ability to represent large data sets, containing many different variables, with a smaller number of composite variables (i.e., components or axes) (Jongman 1995). Those sites, species, or environmental variables associated with the x-axis are those with the most explanatory value due to the statistical and graphing methods of the process. Interpretation of the resultant graph is therefore intuitive in that the location of symbols (species, sites, and/or environmental variables) in proximity to each other and the axes gives indication of correlation and power of interpretation. Tests of significance (e.g., a Monte Carlo permutation test) determine the validity of the axes and the entire model. The interpretation of these significant associations is then vetted based on ecological validity (presented in Section 4).

The statistical assessments performed were indirect and direct ordinations. Indirect ordinations such as Non-metric Multidimensional Scaling (NMS) were used to visualize temporal variance of biological communities (response matrix) at the site level (α diversity) and between sites (β diversity) over time. The use of NMS, when possible, is a purposeful choice over the older PCA (principle components analysis) as it has been constructed to be a more robust ordination procedure related to unbalanced datasets with non-normal distributions (i.e., ecological data) (McCune 2002). Interpretation of the results (graphical) is similar in that the distance between sites (the basis for the interpretation) indicates the similarity or difference of the sampled dataset (e.g., diversity and quantities). The improvements of NMS over PCA relate to the iterative process that the former performs automatically to reduce "stress" (i.e., mismatch between the rank order of distances in the data, and the rank order of distances in the ordination) in the analysis thereby creating a more accurate graphical representation of the true statistical relationship of the sampled community.

Direct ordinations, such as Nonparametric Multiplicative Regression (NPMR), Redundancy Analysis (RDA), and Canonical Correspondence Analysis (CCA), incorporate the same principles of interpretation as the indirect methods above but combine an explanatory matrix in addition to the response matrix. The procedure seeks significant associations between the response matrix (e.g., species), explanatory variables (stressors), and sampling sites (made up of biological data). These results are located on a graph with respect to their association to the stressors and are determined to be significant with respect to the variance in the datasets.

Prior to analyses, all datasets were prepared through a combination of removing rare taxa (found at <5% of sites) and transformations (i.e., relativization by maximum) (Legendre 2001). Unless rare taxa is the focus of an analysis, removing data collected at fewer than 5% of the sampling sites over time is necessary to prevent the statistical ordination from placing emphasis on data erroneously. Similarly, transformation of data reduces the emphasis placed on overly abundant taxa. Relativization, also allows for the equal weighting of environmental parameters, which are often presented with different units of measure. If these two steps were not performed, the analyses would have put excessive weight on rare or over-abundant taxa, which would cloud the overall result and interpretation of ecological reality. Data matrices were combinations of quantitative data (e.g., quantifiable measurements of abundance) and categorical data (e.g., season measured). Specific analysis steps for each group of usable collected data is presented in the subsequent sections. It should be noted that not all provided data were adequately sampled to allow for analysis. Coding of sampling sites on the following graphs consistently follow the format of site abbreviation (e.g., City Park = CP) followed by sampling quarter (e.g. third quarter = 3) and finally year (2013 = 13) for a general appearance of CP313.

Two vegetation datasets were used in the analysis, “Tabular BW Vegetation” and “TWR Physical Observations”. Since the latter was only from three sites on the San Marcos spring system alone, and is a deprecated dataset, emphasis was placed on the “Tabular” dataset since it contains sampling information from both spring systems over a sixteen-year span. The tabular dataset is the numerical expression of the aquatic vegetation GIS mapping efforts. The dataset describes aerial coverage of submerged aquatic vegetation species at the long-term biological monitoring reaches of both systems.

The initial ordination was a NMS (indirect - species only) on the “Tabular” dataset and excluded the Texas wild rice (*Zinzania texana* - excluded since only found at San Marcos). Included taxa were algae, bryophytes, and the genera *Caboma, Certophyllum, Ceratopteris, Chara, Colocasia, Eichhornia, Heteranthera, Hydrilla, Hydrocotyle, Hygrophila, Justicia, Limnophila, Ludwigia, Myriophyllum, Nasturtium, Nuphar, Potamogeton, Rorippa, Sagittaria, Vallisneria,* and *Zizaniopsis.* The matrix was constructed by site and date to visualize those potential sources of variation. The NMS ordination was performed within the PC-ORD statistical software package (McCune 2016). Separate NMS analyses focused on differences between sites and over time within each spring system.

The “TWR Physical Observations” dataset was formatted for the direct ordination procedure NPMR and analyzed with the statistical package Hyperniche (McCune 2009). This data describes the overall health and environmental conditions of designated vulnerable stands of Texas wild-rice. The yearly “area” measure for Texas wild rice was the sole response variable with estimates of predation, percent emergent vegetation, and flow as explanatory variables. It should be noted that the results of this survey and the resultant analysis should be interpreted with caution due to issues with consistent representative data as indicated by the EAA.

The final vegetation analysis was on the non-wild rice species surveyed for area extent and compiled within the Tabular BW Vegetation spreadsheet. The direct ordination procedure of CCA was performed using the vegetation area as the response variable and discharge, preceding low or high flows, and season as explanatory variables.

2.1 Fish

With emphasis on the fountain darter (*Etheostoma fonticola*), four sets of count data were used for fish analysis: “Fountain Darter Visual Observations”, “Drop Net”, “Timed Dip”, and “Dip Net”. The “Fountain Darter Visual Observations” dataset featured results from Landa Lake from 2003-2015. The “Drop” and “Dip” surveys were from both rivers with site-specific vegetation and flow from 2000-2015.

The first analysis of the “Fountain Darter Visual Observations” dataset was a simple correlation plot (Excel) of darter numbers versus site flow. Secondly, the “Timed Dip” was analyzed with NPMR to determine a potential effect of seasonality and location. Next, an NMS was performed on a single matrix of “Random Dip” data for which darter numbers were plotted per vegetation type and then were analyzed per site to determine vegetation preference. “Fixed Dip” was then analyzed for the effect of substrate type on darter abundance within NMS. Finally, the “Drop Net” fish dataset was analyzed with NPMR in an attempt to build an explanatory model that described the fountain darter site variation based on the differences of “dominant vegetation”, “vegetation coverage”, “substrate”, “flow”, “water depth”, “temperature”, dissolved oxygen”, “specific conductivity”, and “pH”. This data set was also split by river/spring system for two more individual NPMR tests.

2.2 Macroinvertebrates

Three datasets for macroinvertebrates were selected for analysis. This biological group was sampled both to monitor the status of the endangered riffle (*Heterelmis comalensis*) and dryopid (*Stygoparnus comalensis*) beetles and to determine potential food sources for the fountain darter. “Macro drift” data represents passive sampling with a net and bucket placed over four spring sources from 2003-2015. The “macroinvertebrate” dataset came from seven sites (Comal and San Marcos springs) from 2013-2015. Lastly, the “lure” dataset comes from three sites (ten total springs from 2004-2015) sampled with a cotton “lure” inserted into the spring orifice for four weeks to ensure adequate colonization.

Due to substantial differences in the macroinvertebrate data from some of the sites/dates in the “macro drift” and “macroinvertebrate” datasets, outlier removal was initially required to produce a valid NMS analysis. Outlier identification analysis was performed within PC-ORD and the result identified 4-5 site/dates that required removal. Following this step, the “macro drift” and “macroinvertebrate” datasets were analyzed with NMS to discern spatial and temporal variation among sites. Additionally, a direct RDA analysis was performed within the software package CANOCO (ter Brakk 2014) on the “Macro drift” data with only the endangered species as response variables (i.e., *Heterelmis comalensis, Stygoparnus comalensis, and Stygobromus pecki*) and temperature, conductivity, pH, dissolved oxygen, velocity, and sampling quarter as explanatory variables. Finally, the “lure” dataset was graphed (Excel) with average flow to discern trends during the sampling period.

2.3 Salamanders

Only one data set existed for analysis (“salamander”) based on diving observations from both springs. The date range was from 2002-2015. The NPMR analysis was performed on this data using corresponding water quality and potential macroinvertebrate prey data.

2.4 Water Quality

Not only were water quality results used as explanatory variables in the analysis for the species of concern, significant effort was also put solely into the analysis of the water quality results between sites and over time. Due to consistent and extensive sampling, the focus of this analysis was the "BW Water Quality" data matrix. This data was collected from 2000-2015 at up to 35 sites within both spring systems. The results of the analysis consist of physiochemical in situ measures (i.e., pH, conductivity, temperature, and dissolved oxygen), total suspended solids, and several constituents of phosphorus and nitrogen. One important aspect to consider with this extensive dataset is the impetus for sampling: critical flows. Since "normal" flow sampling historically resulted in very little variation in water quality, the monitoring impetus switched to excessive flooding or low flow events. The diluting or concentrating effects of the extreme flows were deemed the most important periods of concern regarding water quality stressors on biological populations. As with the biological results, the initial analysis for this data was a NMS to visibly determine patterns in the data related to spatial and temporal variability. Second, a RDA was performed using temperature and season as explanatory variables.

2.5 Recreation

A brief attempt to summarize the recreational pressure on the spring systems was made by graphing the measures of in-water recreation (i.e., tubing and swimming) by site, year, and quarter. This dataset was collected by Texas Master Naturalists.

3 RESULTS

3.1 Vegetation

Vegetation coverage (area) was found to vary substantially with regards to spatial and temporal measures and to some extent to flow (based on significant correlations). As demonstrated by the NMS in Figure 3.1, spatial variation was obvious as distinct groupings among sampling sites was prevalent. Prior to the analysis, *Zizania* was removed since it is absent from the Comal Springs systems and would therefore bias the analysis for known reasons. Landa Lake (LL), City Park (CP), Old Channel (OC) and Spring Lake Dam (SL) had the most consistently exclusive communities and area coverage over time. The sites I-35 (I3), Upper Spring Run (US), and Lower New Channel (LN) sites, however, varied more over time. Landa Lake and Upper Spring Run consistently had the most diversity and greatest coverage of aquatic plants over time (e.g., most abundant coverage of *Nuphar, Ludwigia*, algae, *Cabomba, Vallisneria, Sagittaria*, and bryophytes) whereas the Old Channel reach had the lowest abundance of plants over time (*Ceratopteris* was the only taxa that was consistently most abundant at the Old Channel). Of note is the high "stress" of the analysis (stress result of 13 = limited explanatory power) due to the large dataset. "Stress" is a value measured in the NMS to determine how much explanatory power can be given to the analysis result. High amounts of stress indicate excessive site-level variance in datasets and limit the valid interpretation of the result. Low stress indicates low site-level variance and greater confidence in the analysis result.

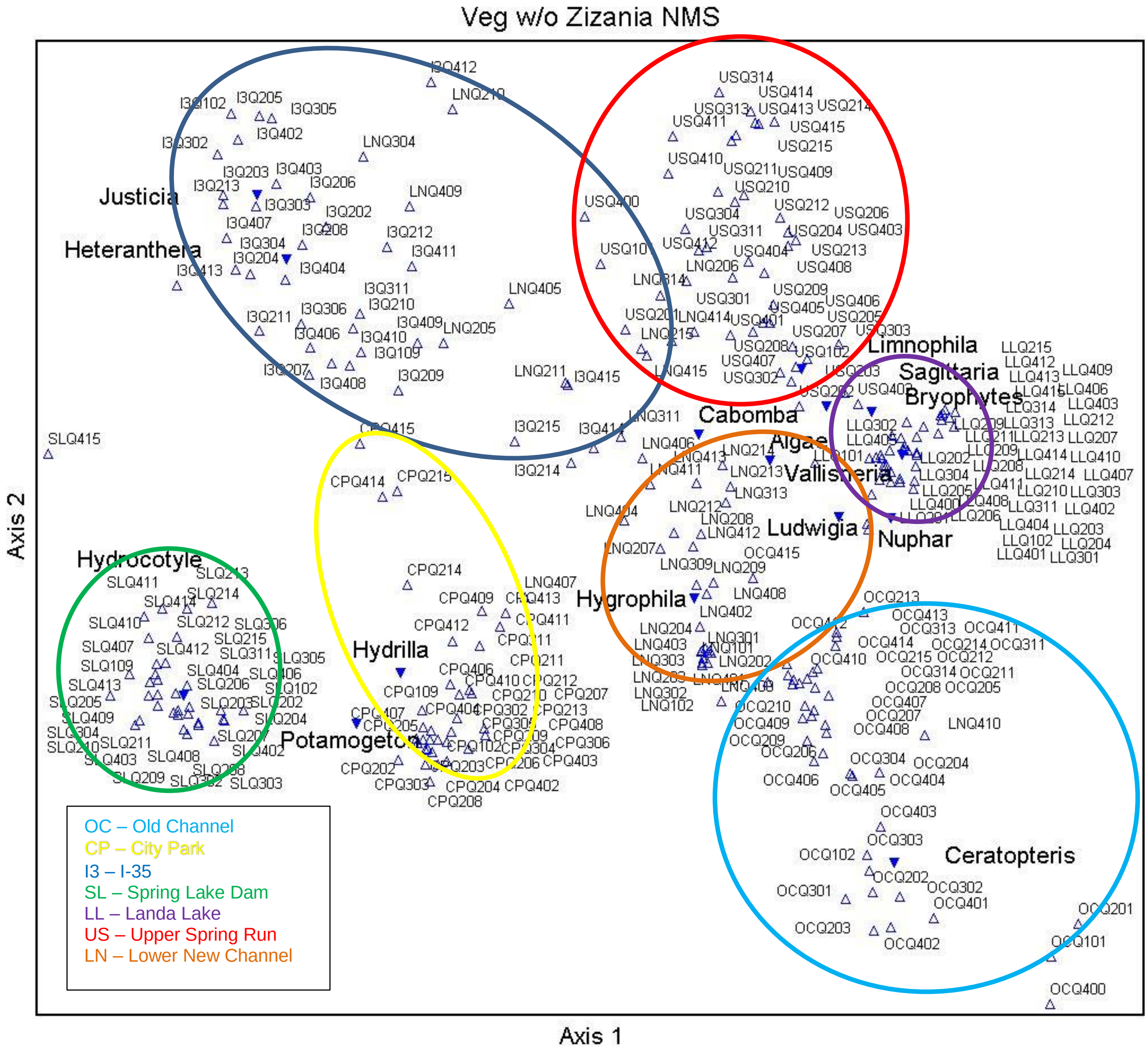

Figure 3.1. Non-metric Multidimensional Scaling (NMS) triplot of sites based on diversity and area of vegetation at Comal and San Marcos springs (Tabular Dataset – without *Zizania*). Circles indicate the most-dense grouping of individual sites based on similarities of vegetation coverage between sites (spatial) and time (temporal).

Error! Reference source not found. is a NMS graph of the Comal Springs portion of the dataset only that sought to uncover the pattern of vegetation diversity/coverage between the long-term sampling sites in that system. Landa Lake exhibited the smallest amount of temporal variation between sampling dates followed by Lower New Channel. Upper Spring Run and the Old Channel exhibited much larger temporal variation indicated by the spread of site/date symbols on the graph. Landa Lake also had the most abundant and diverse species assemblage as it was associated with algae, *Nuphar, Ludwigia, Vallisneria, Sagittaria*, and Bryophytes.

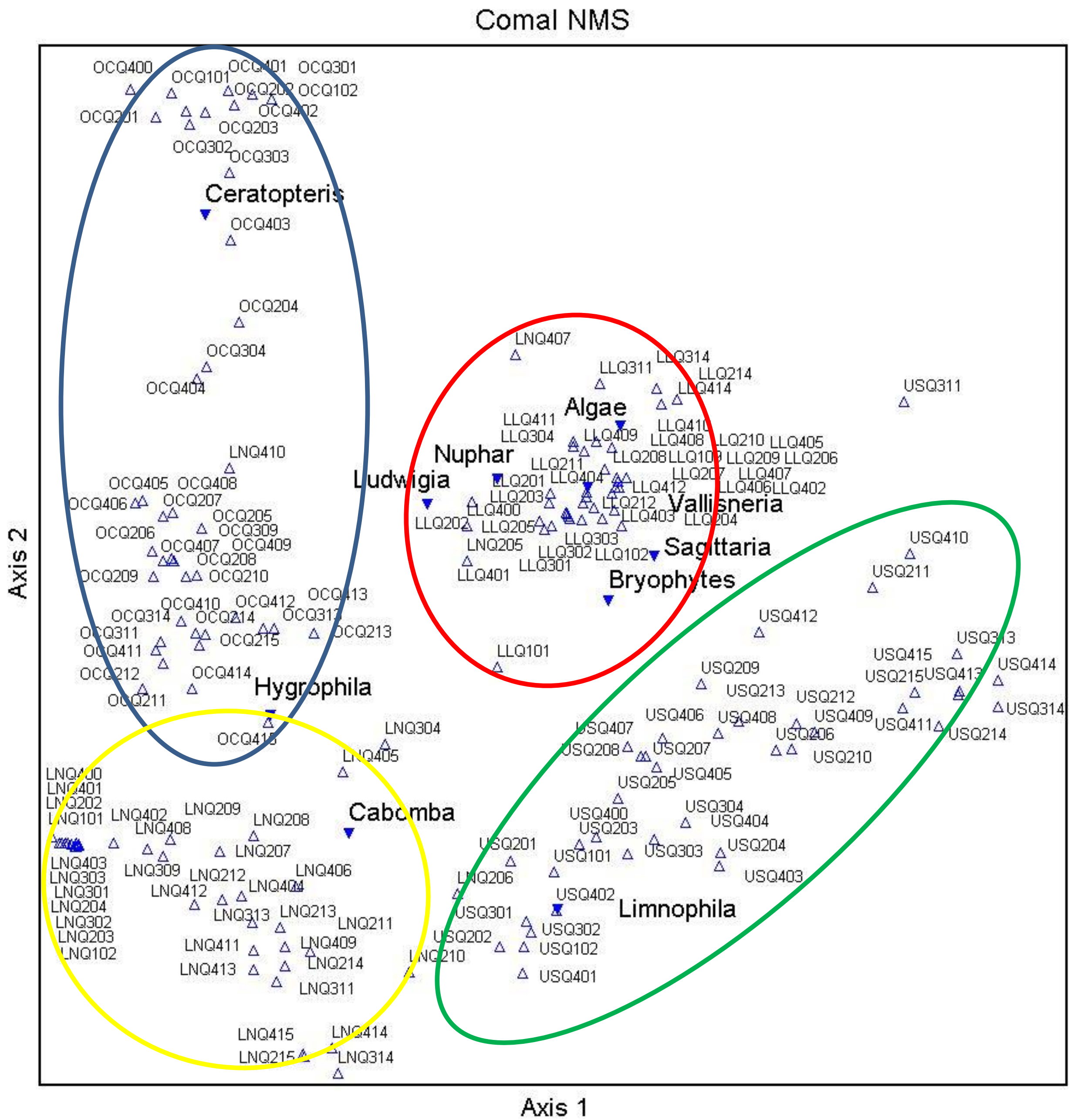

Figure 3.2. NMS scatterplot of sites based on vegetation differences at Comal Springs. Landa Lake = LL, Upper Spring Run = US, Lower New Channel = LN, Old Channel = OC.

The subsequent NMS of the vegetation at the San Marcos springs was similar in its site (spatial) separation indicating distinct differences between sampling reaches. The NMS in Figure 3.3 indicates the least abundance and diversity of the measured plants at Spring Lake Dam (SL) with only two species, *Vallisneria* and *Hydrocotyle*, having the greatest abundance there. The I-35 crossing of the San Marcos River (I3) and City Park (CP) had the greatest abundance and diversity with Bryophytes, *Ludwigia*, *Cabomba*, *Justicia*, and *Heteranthera* dominating at I-35 and *Hygrophila*, *Potamogeton*, *Hydrilla*, *Sagittaria*, and algae dominating at City Park.

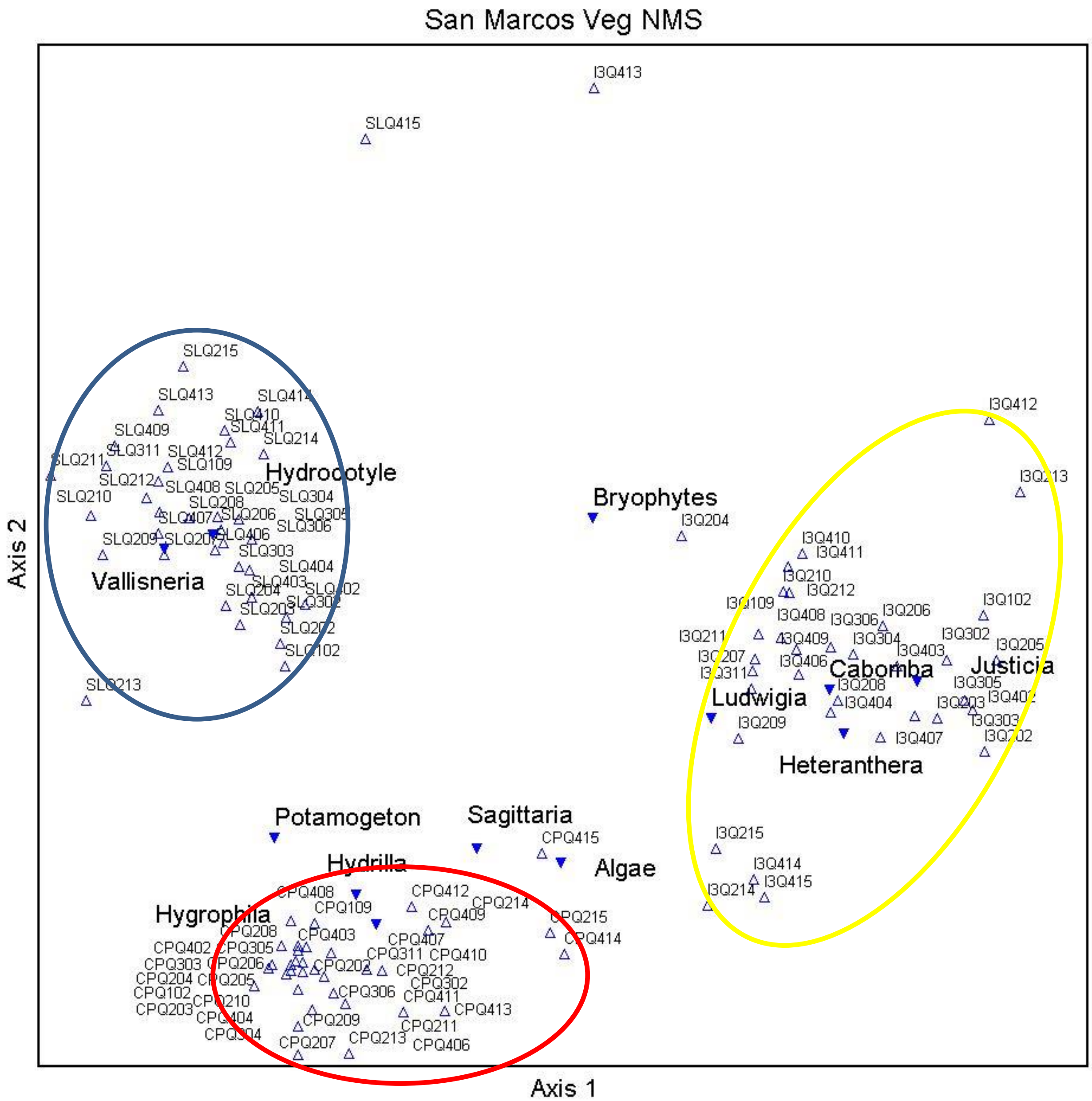

Figure 3.3. NMS graph of San Marcos vegetation as it differs by site. City Park = CP, Spring Lake Dam = SL, I-35 = I3.

In an attempt to determine the source of the aquatic vegetation variance, a CCA was performed with both of the spring's survey information (response variables) and the limited environmental information (explanatory variables). As the primary expected stressor, flow was scrutinized from adjacent USGS real-time gages to provide a sample date monthly discharge average and its corresponding indication of flow type (low, normal, or high as indicated by the EAA: Comal low-flow ≤ 200 cfs, high-flow ≥ 500 cfs; San Marcos low-flow ≤ 120 cfs, high-flow ≥ 385 cfs; normal flow is middle flow value for each system, respectively), as well as the number of either low flow or high flow events since the previous vegetation survey. In addition to flow, the season of sampling was included as a categorical variable to determine a potential seasonal variation in the communities. As demonstrated in Figure 3.4, three of the flow variables were found to have significant correlations with variance in the vegetation data: monthly discharge average (Discharge), Flow Category (low, normal, or high), and the number of preceding high flow events (High Flow Events). This model was significant ($P<0.002$) however it only explained 9% of the vegetation variance. The season of vegetation sampling was not found to be significant.

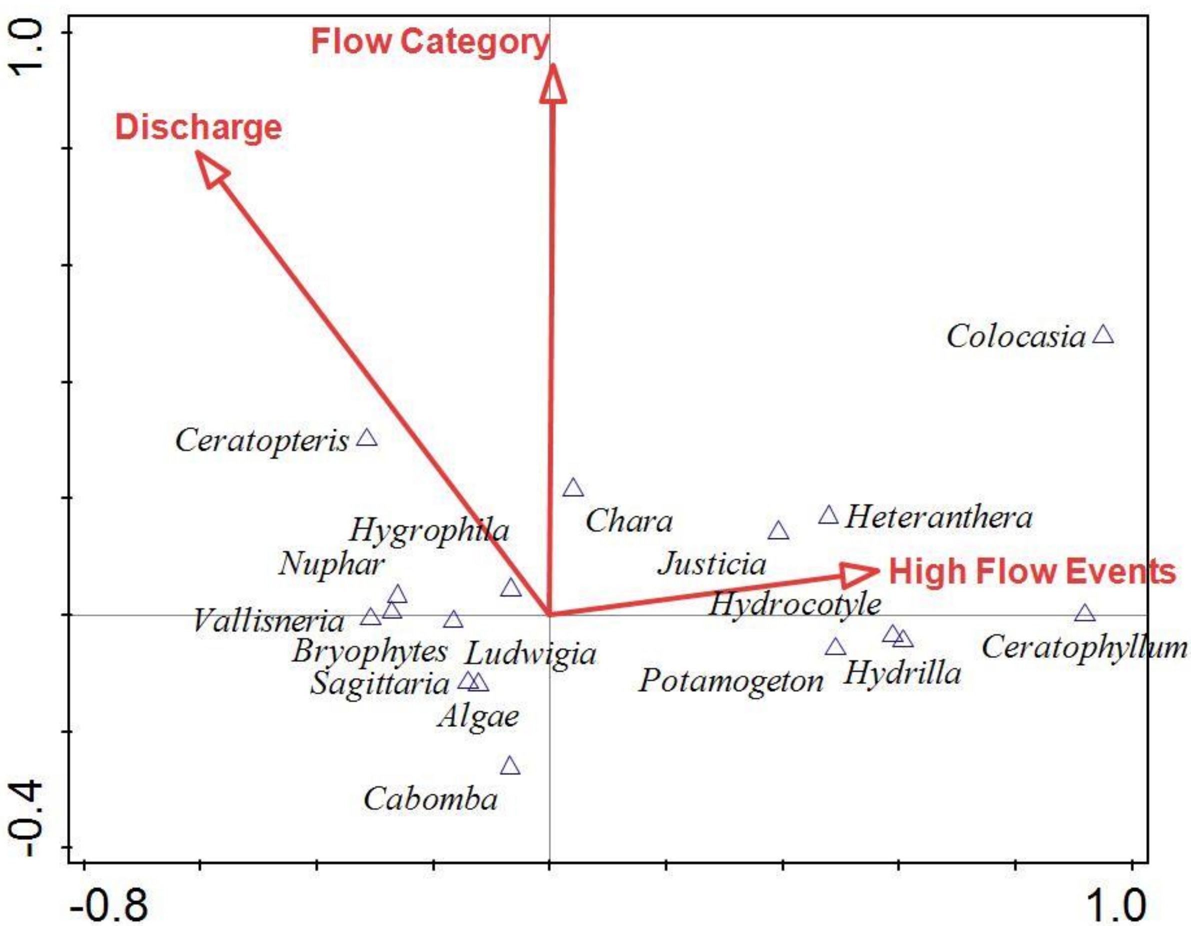

Figure 3.4. Canonical Correspondence Analysis (CCA) of vegetation area in response to flow (Comal and San Marcos springs).

Using the “TWR Physical Observations” dataset, Texas wild rice was scrutinized for patterns of variance due to stressors by itself. Figure 3.5 and Figure 3.6 are three dimensional graphs of the same NPMR test, which seeks to develop models of association by matching correlations of variance between the plants and the measurement environmental stressors. The measured stressors were average flow, visible evidence of predation, and percentage of emergent vegetation (shading effects). The model with percent emergent plants and predation (Figure 3.5) had an r^2 of 0.35 while the model with predation and flow (Figure 3.6) had an r^2 of 0.48, both of which indicate a fair correlation for ecological analyses.

On Figure 3.5, a trend exists between decreasing vegetation area with increasing coverage of emergent vegetation as well as a slight increase in vegetation area with increasing predation. Figure 3.6 alludes to a stronger positive relationship between vegetation area and predation. The negative effects of high flow are also indicated on this graph.

As previously mentioned these results should be interpreted with caution as the measures of the stressors are not believed to be rigorous and consistent over time.

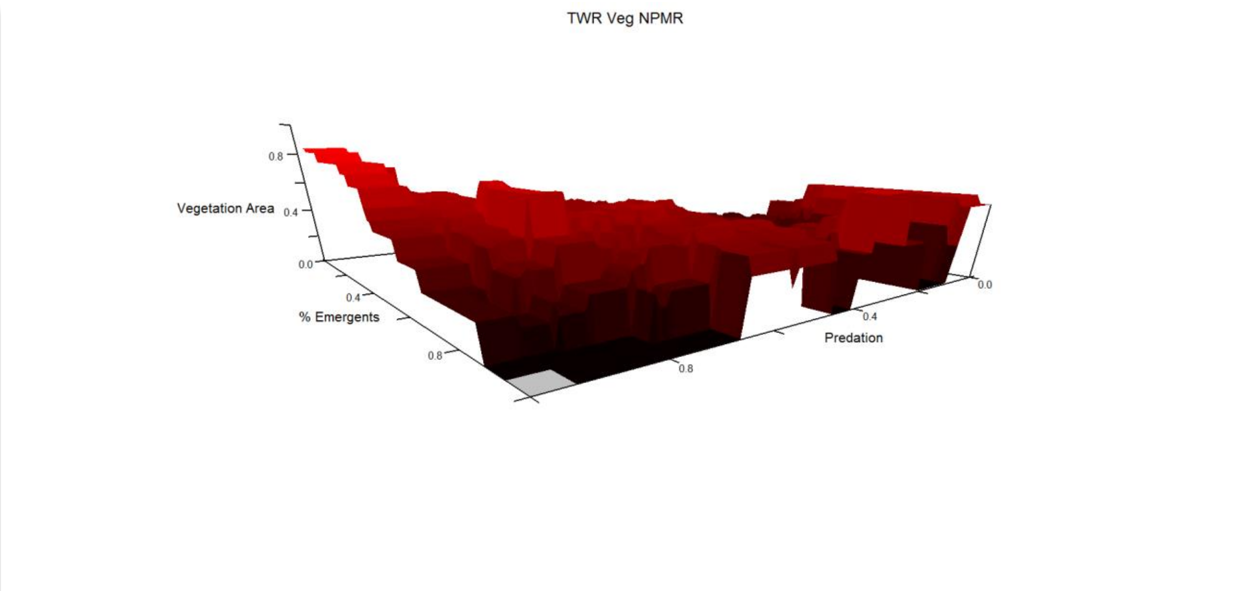

Figure 3.5. Nonparametric Multiplicative Regression (NPMR) graph on wild rice area at San Marcos spring versus predation and emergent vegetation abundance (r^2 = 0.35).

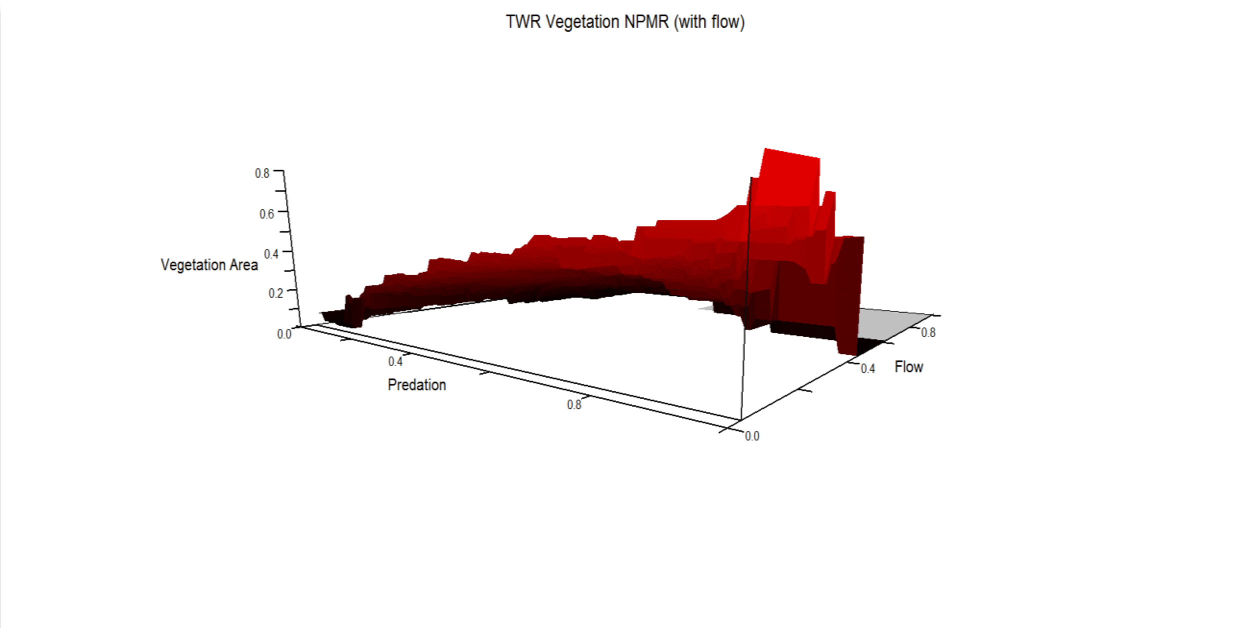

Figure 3.6. NPMR graph on wild rice at San Marcos spring versus predation and flow (r^2 = 0.48).

3.2 Fish

From the “Fountain Darter Visual Observations” dataset, the initial simple correlation graph of darters abundance and vegetation coverage eludes to a positive relationship as the trend of both matches with close similarity (See Figure 3.7). The number of fountain darters observed was normalized to the maximum number of fountain darters observed; as such, both variables could be plotted on the same y-axis as a scale from 0 to 100 percent.

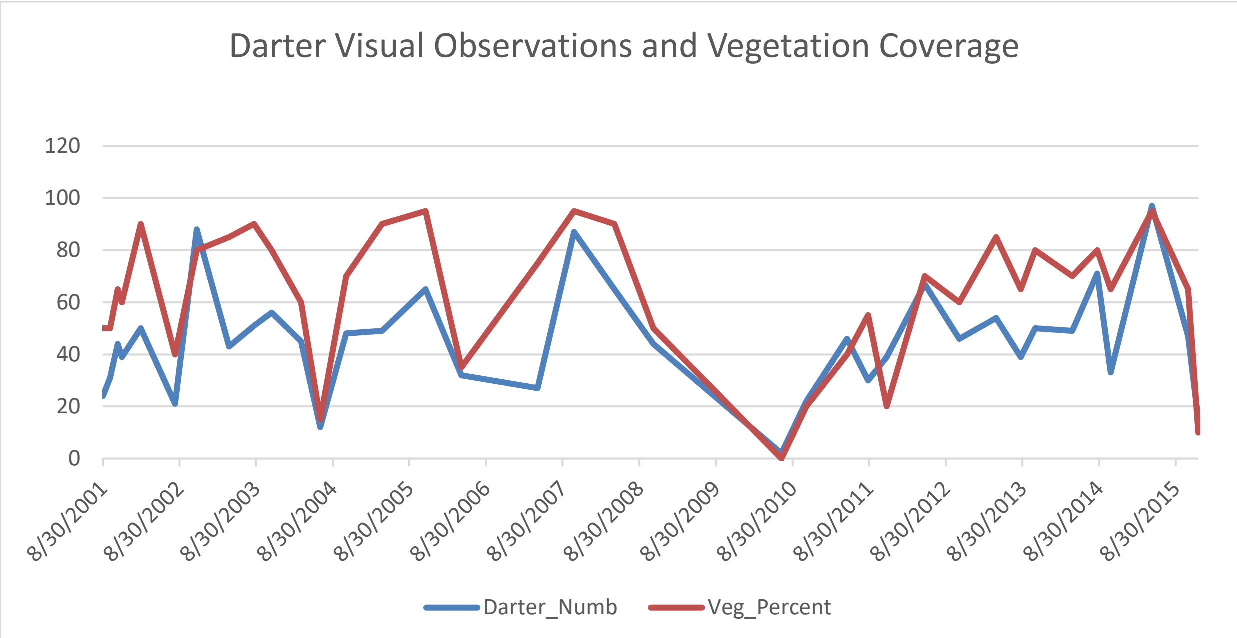

Figure 3.7. Relationship between the percentage of vegetation coverage in plot versus the number of fountain darters observed.

More specific species/vegetation affinities were observed by the NMS of darter abundance by vegetation type from the “Random Dip” dataset. Figure 3.8 is the result of an NMS with a different type of site coding of the analysis matrix. From both spring systems, the abundance of darters per vegetation type was coded by site. Therefore, the lengths of the blue arrows on the graph (with the vegetation label) indicate the abundance of darters in relation to particular sites. In other words, the graph indicates which plants are associated with the greatest abundance of darters at each site. As with all the ordination triplots, the arrows/sites associated with the x-axis have the greatest explanatory power. With this information in mind, at Spring Lake and New Channel the greatest abundances of fountain darters were found in greatest association with *Cabomba* (most significant), *Sagittaria*, and *Milfoil*. At Spring Lake Dam, City Park, and at I-35 several vegetation species shared the greatest abundance of darters. *Potamogeton*, or mixes of this genera with others, was the most supportive species, yet *Vallisneria*, *Hydrilla*, and *Hydrocotyle* also provided ample habitat for the fish. Lastly, Upper Spring Run, Landa Lake, and the Old Channel sampling reaches had the greatest abundance of darters in conjunction with *Ludwigia*, algae, and bryophytes.

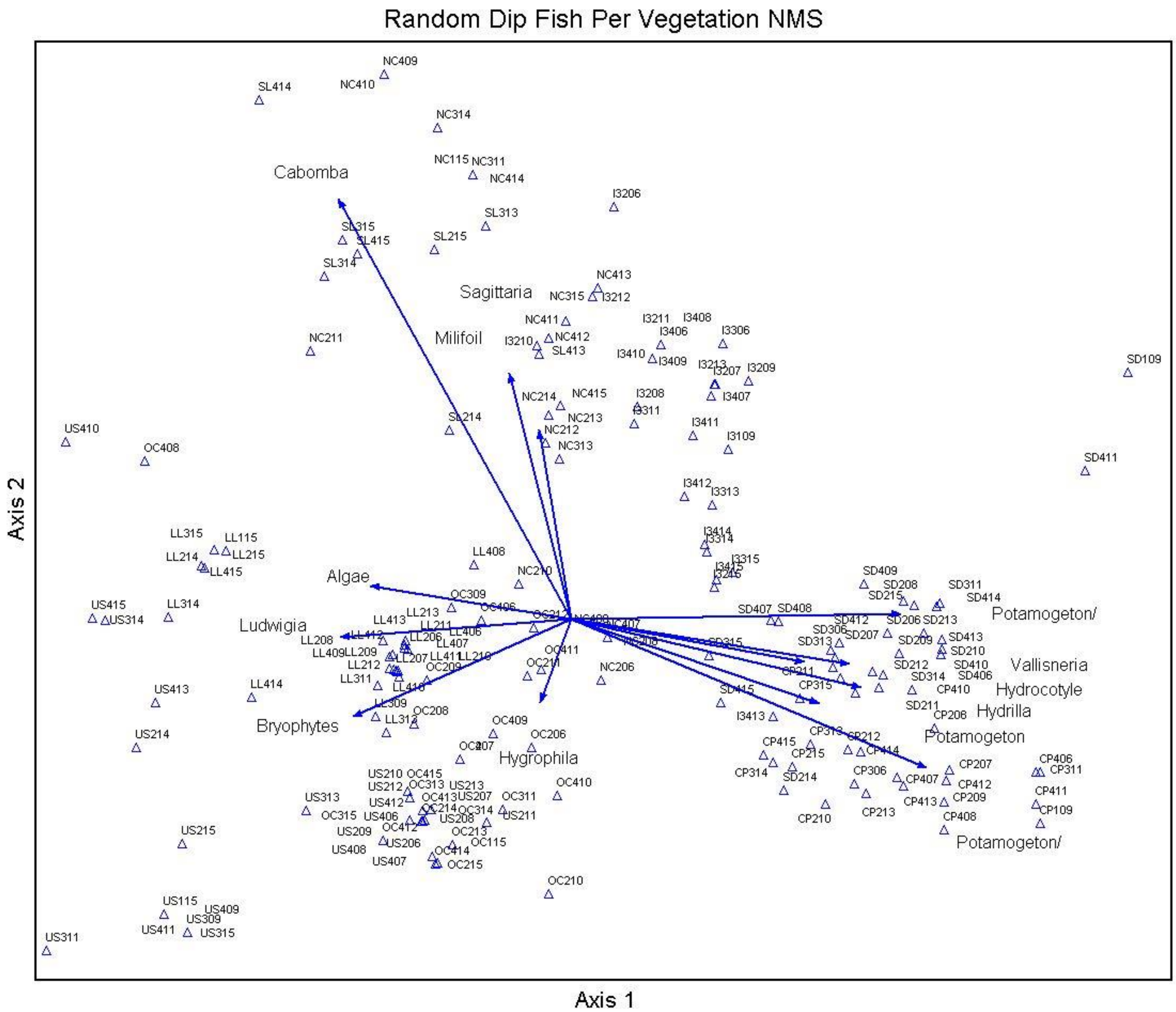

Figure 3.8. NMS ordination graph of fish abundances per vegetation affinities in relation to sites at Comal and San Marcos springs.

A simple bar chart of darter abundance from the "Random Dip" dataset (Figure 3.9) indicated the greatest abundance of darters from Landa Lake particularly in association with Bryophytes. Spring Lake and Upper Spring Run had the lowest darter abundance. A visual comparison between sites does not reveal a particular plant preference with regard to its effect on darter abundance.

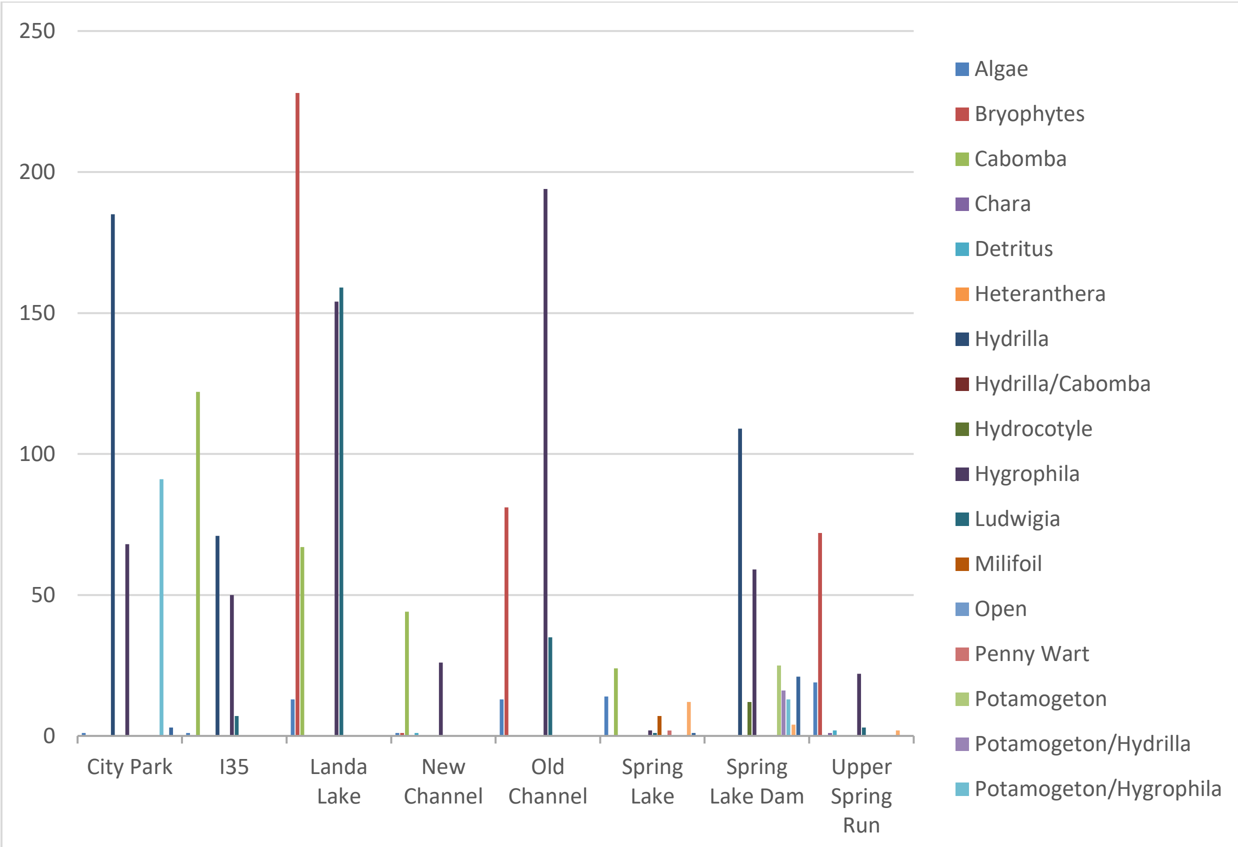

Figure 3.9. Site abundances of fountain darters which indication of plant preference.

The final analysis on the fountain darter was the NPMR to determine the overall significance of vegetation type, vegetation coverage, substrate type, flow, water depth, temperature, dissolved oxygen, specific conductivity, and pH. The NPMR analysis indicated that too much information clouded the results; however, it did implicate temperature, the abundance of bryophytes, and overall vegetation cover as significant models to predict variation in species abundance by site. This result should be interpreted with caution as the r^2 was only 0.2. This is the reason for the exclusion of the resultant graph as discernable trends were not clearly apparent.

Subsequent to the holistic analysis above, two additional NPMR tests were performed on each spring system individually. However, the Comal Springs analysis was almost identical to the overall NPMR analysis described above, and the San Marcos analysis did not yield any significant models.

3.3 Macroinvertebrates

The results of the threatened and endangered macroinvertebrate NMS (“macro drift” dataset) did not reveal a significant model indicating a lack of discernable variance over time and between sites for those three taxa (i.e., *Heterelmis, Stygoparnus,* and *Stygobromus*). However, the RDA analysis between the threatened and endangered taxa and the measured environmental variables for the same dataset (Figure 3.10) did yield a significant direct ordination model ($P<0.05$) that indicated a pattern of temperature influence (21% of the measured community variance) with a lessened correlation of sampling season (6% of variance). Based on the associations seen in Figure 3.8, *Stygobromus* individuals were found in greatest numbers in sites with higher temperature (i.e., Western Upwelling), whereas *Heterelmis* and *Stygoparnus* exhibited the opposite preference to sites with the lowest temperatures (Spring Run 1 and 3). Even though they differed in temperature preference, it is worth noting that velocity did appear to have a positive influence on both *Stygoparnus* and *Stygobromus* as those two taxa's abundance trended toward the higher velocity sites.

The last macroinvertebrate analysis focused specifically on *Heterelmis* and the correlation in its abundance with flow. As can be seen in the correlation plot in Figure 3.11, the trend flow average and numbers of riffle beetles appear to have a positive relationship.

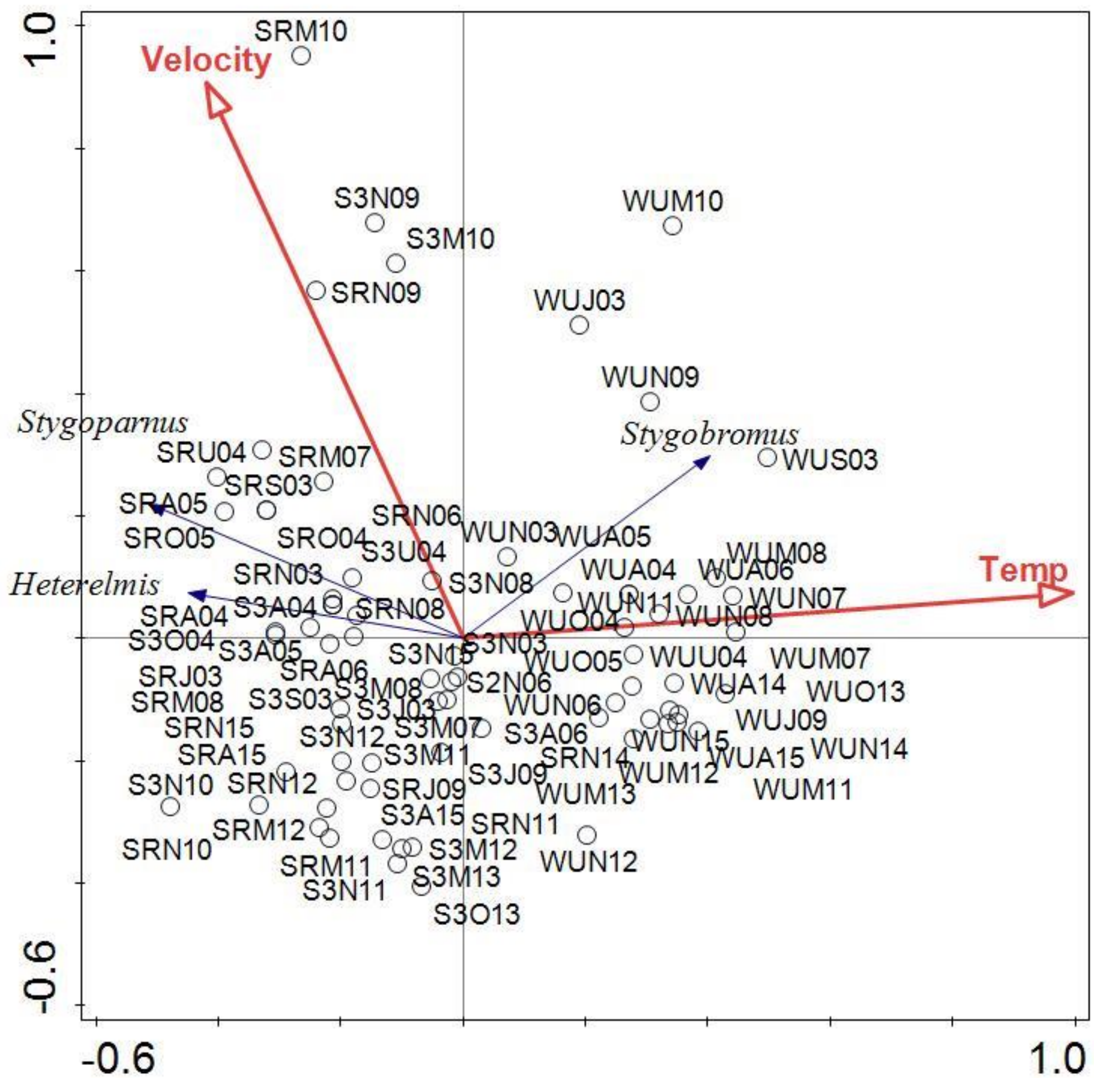

Figure 3.10. Redundancy Analysis (RDA) of *Heterelmis, Stygoparnus,* and *Stygobromus* survey results from Spring Run 1 (SR), Spring Run 3 (S3), and Western Upwelling (WU) as compared to measured environmental variables.

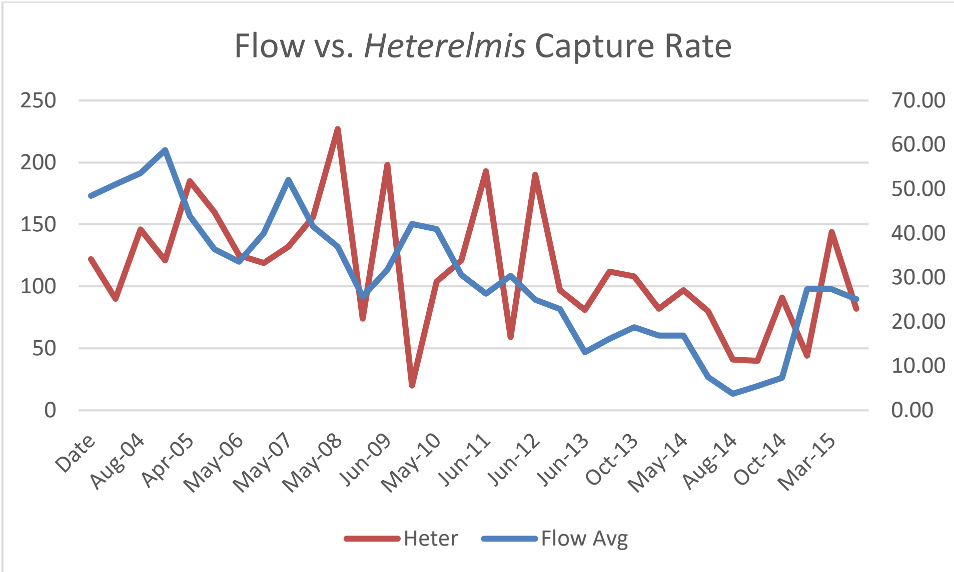

Figure 3.11. Flow rate daily average plotted against sampled *Heterelmis* individuals (cotton lure) from the same date. The left vertical axis is *Heterelmis* individuals and the right vertical axis is flow rate in cfs.

3.4 Salamanders

The resultant NPMR from the one multivariate salamander analysis presented interesting trends between water quality, temperature, and prey availability (i.e., macroinvertebrates taken from the similar site/dates from the "macro drift" dataset). Figure 3.12, Figure 3.13, and Figure 3.14 show the results of the same NPMR only presented with various combinations of the model's significantly associated variables. The salamander abundance varied positively with the macroinvertebrates *Seborgia* and *Hydracarina* (mites) as seen in Figure 3.12 indicating a potential prey prevalence. However, the reduced R^2 value of 0.3 indicates that caution should be used in this interpretation. Figure 3.13 shows the association between salamander abundance and the water quality parameters of NO_3 and dissolved oxygen alternatively exhibited a very high R^2 of 0.91. This result indicates a strong positive association between salamanders and dissolved oxygen and a negative association with NO_3, however this strong trend is likely an artifact of the sampling protocol. Figure 3.14 indicates that increasing temperature also overrides high dissolved oxygen as temperature had a negative association with the salamander abundance (R^2 = 0.90). Figure 3.15 indicates that temperature and NO_3 both have negative associations with the salamander abundance (R^2 = 0.90). Figure 3.16 is a correlation plot between salamander abundance and NO_3 concentration from sites at both spring systems. This correlation implies a strong negative response of the salamanders to nitrates.

Further vetting of these trends resulted in an alternative explanation that implies that "correlation doesn't always equal causation". It is probable that the significant correlations described above are likely an artifact of the sampling protocol. Water quality samples were only collected during "critical flow" periods – when either flows were very low or very high. Both water quality parameters and salamander numbers (either a result of flow impacts or due to alterations of sampling efficacy) vary with significant changes in flow. The correlations described above are likely coincidental and not cause-and-effect.

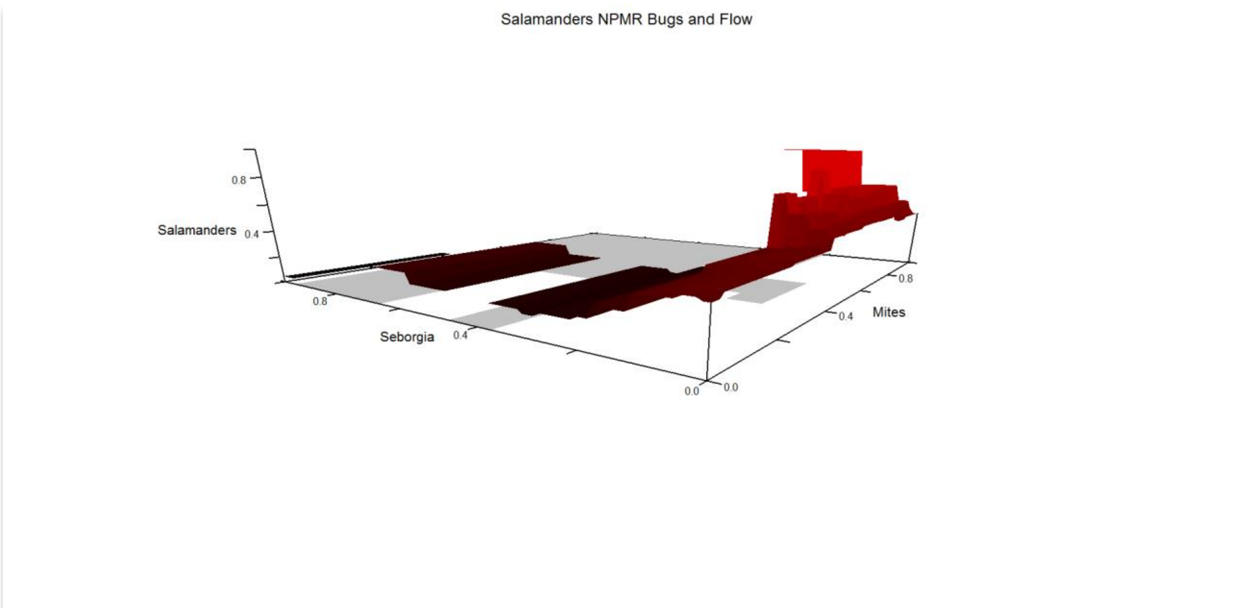

Figure 3.12. NPMR model graph of salamander observations and significantly associated abundance of "prey" (i.e., *Seborgia* and water mites). R^2 = 0.30.

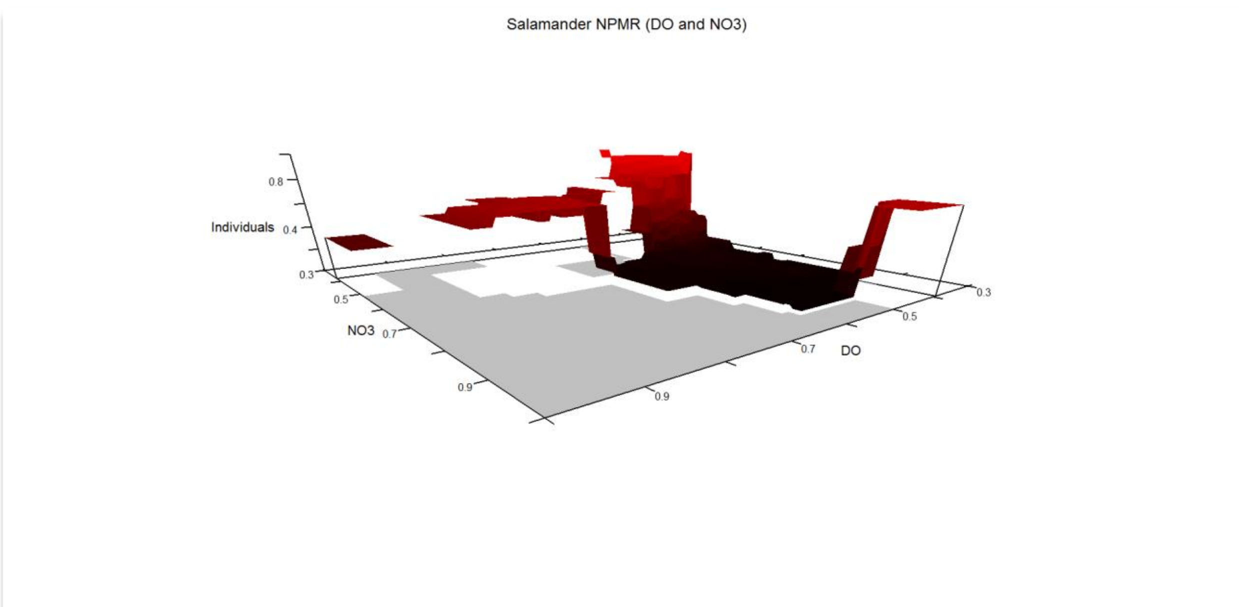

Figure 3.13. NPMR model graph of salamander observations and significantly associated water quality results (i.e., NO_3 and dissolved oxygen). R^2 = 0.91.

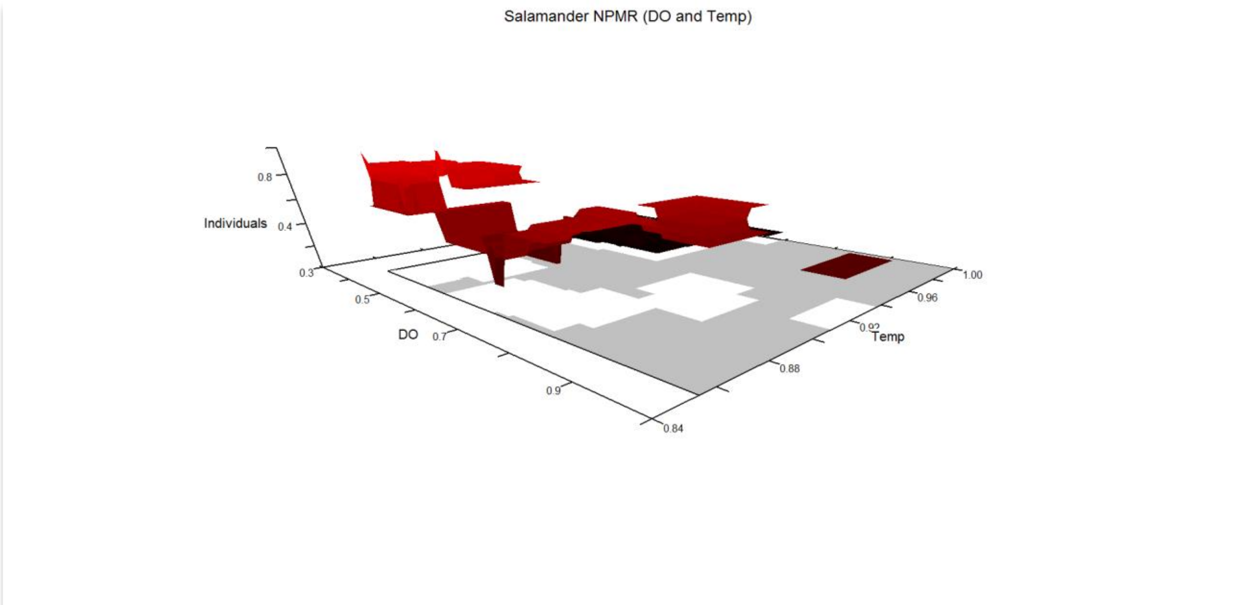

Figure 3.14. NPMR model graph of salamander observations and significantly associated water quality results (i.e., dissolved oxygen and temperature). $R^2 = 0.90$.

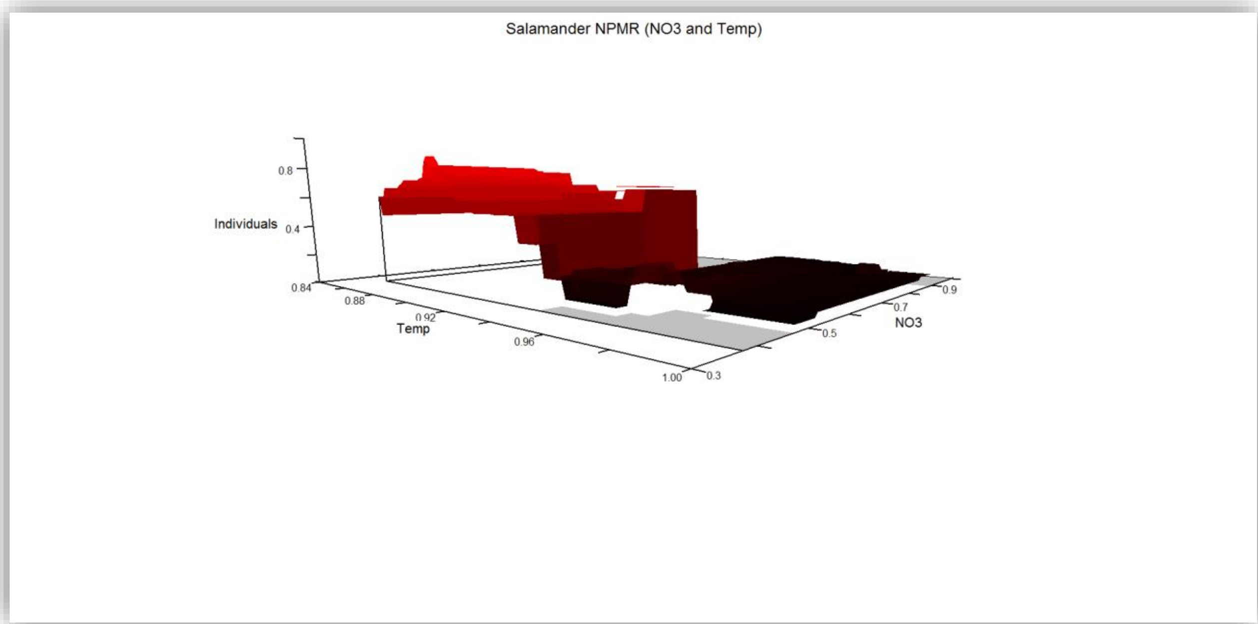

Figure 3.15. NPMR model graph of salamander observations and significantly associated water quality results (i.e., temperature and NO_3). $R^2 = 0.90$.

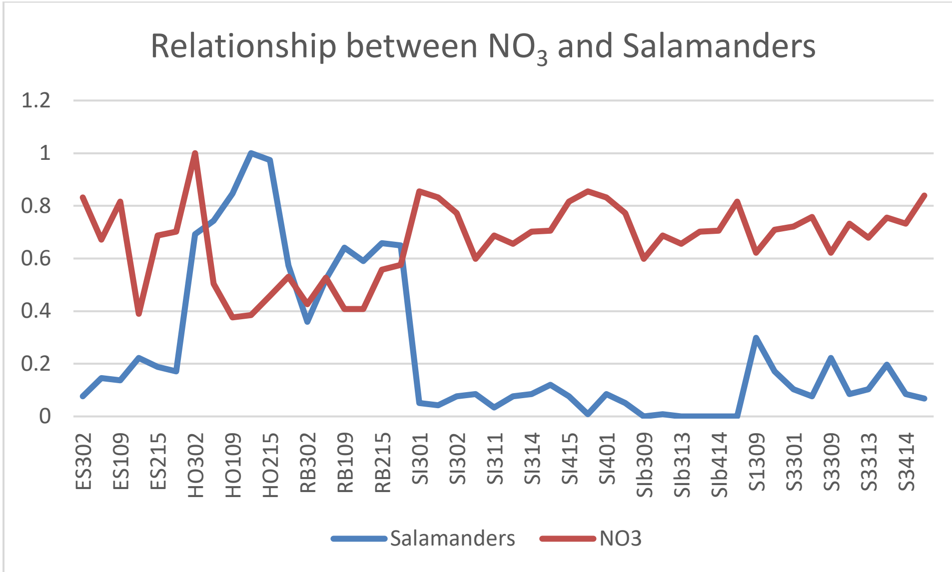

Figure 3.16. Salamander relationship with NO3 concentrations from both spring systems. Data is relativized by maximum.

3.5 Water Quality

The NMS of all the measured water quality parameters from the "BW Water Quality Revised" dataset resulted in a distinction in the results from 2014 (Figure 3.17). Most of the sites/dates exhibited a random location (i.e., no temporal or spatial distinction) on the ordination plot, yet the August 2014 data from Spring Run 1, 2, and 3, Old Channel, Upper Spring Run, Lower New Channel, Island Park Near Channel, Island Park Far Channel, Island Park, Heidelberg Main Channel, New Channel Upstream, and Union Avenue exhibited a significant distinction from all other sites/dates.

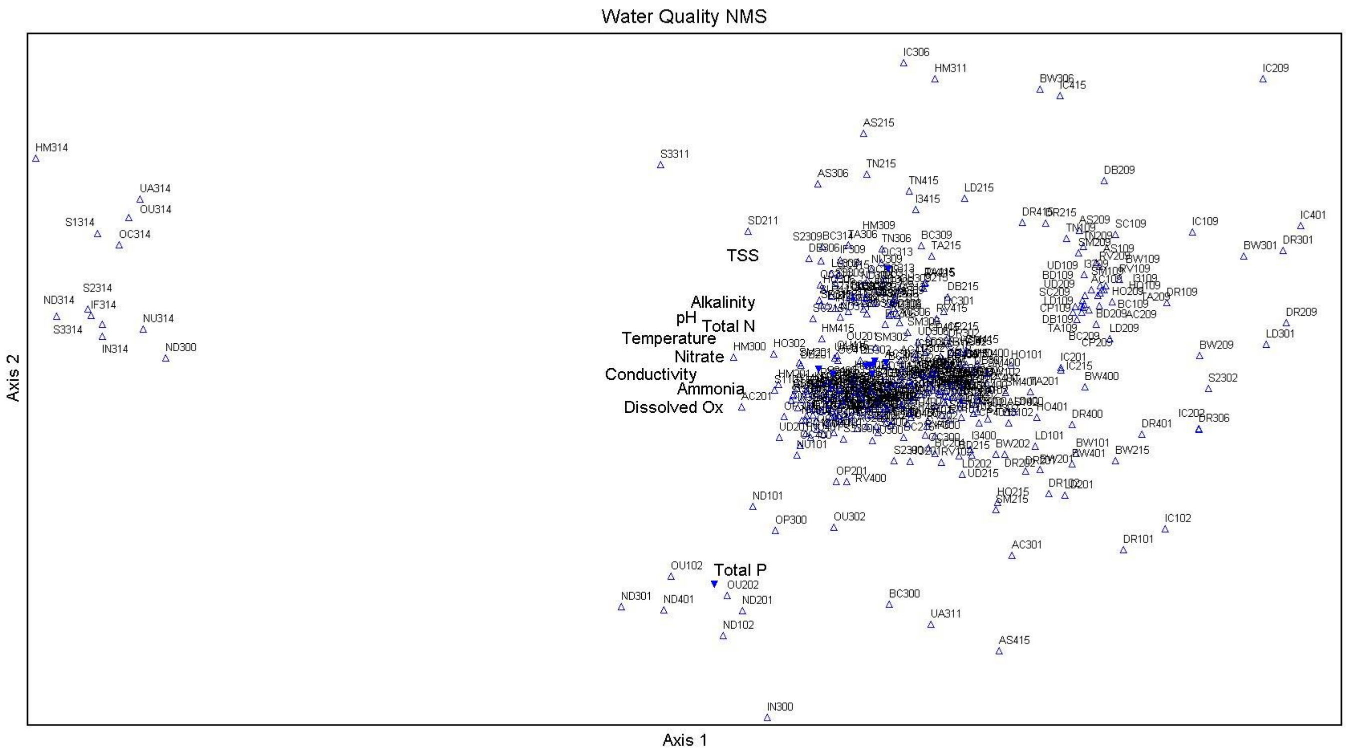

Figure 3.17. NMS diagram of "BW water Quality" dataset. Plot indicates outliers at many sites during August 2014.

The two water quality parameters with the highest correlation of site level variance were found to be Nitrate and Total Nitrogen (Total N). When viewed as vectors (arrows) these parameters point in the direction of the outlier sites from August 2014 indicating that these are the primary parameters setting these sites/dates apart from the rest of the samples. Most of the other parameters were also found at their highest concentrations at these sites (i.e., alkalinity, conductivity, pH, and temperature). It should be noted that the measured dimensionality "stress" of this NMS was relatively high (15.9) (McCune 2002) indicating that caution should be taken when interpreting these results. It's probable that this high stress was due to the large data set though we believe the variation in the 2014 results is valid based on interpretation of the results especially given that 2014 has been indicted as a drought year and nitrate/nitrogen would be expected to become more concentrated in low flows.

In an attempt to understand the significance of water quality variance, a direct RDA ordination was performed on the same dataset while constraining the water quality parameters by temperature and sampling quarter. A robust, significant ($P<0.05$) model resulted (Figure 3.18) that had the greatest variance explained by temperature alone (35%). Sites with the highest temperature had the greatest overall concentrations of ammonia, total nitrogen, nitrate, alkalinity, conductivity, pH, and dissolved oxygen. Coincidentally, the August 2014 sites were again indicated as having the highest levels of the same parameters indicated in the initial NMS (i.e., total nitrogen, nitrates, alkalinity, pH, temperature, and dissolved oxygen). These sites included Island Park Near Channel, Old Channel Downstream, Old Channel Upstream, New Channel

Upstream/Downstream, all Spring Run sites, and Heidelberg Main Channel. Temporal explanatory variables were found to be significant (Quarter 2 and 3) but were only attributed with 2% of the variance in the measured parameters.

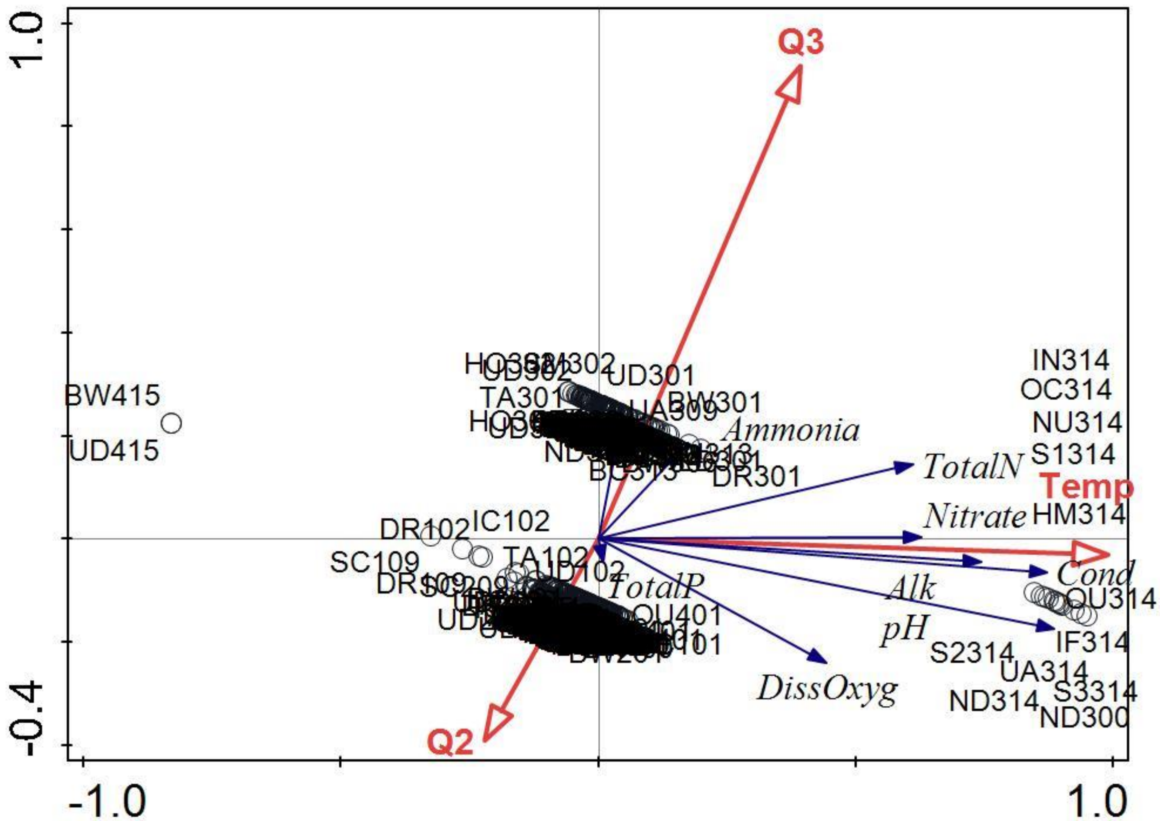

Figure 3.18. Redundancy Analysis (RDA) of water quality results constrained by temperature and quarter sampled.

3.6 Recreation

The summary graph of recreational pressure indicated that the highest recreational pressure appeared to occur consistently in the New Channel and Union Avenue especially during quarter three. Recreation also appears to have increased over time at these two sites. These two sites were also indicated as being outliers due to excessive nitrates and total nitrogen during the third quarter of 2014.

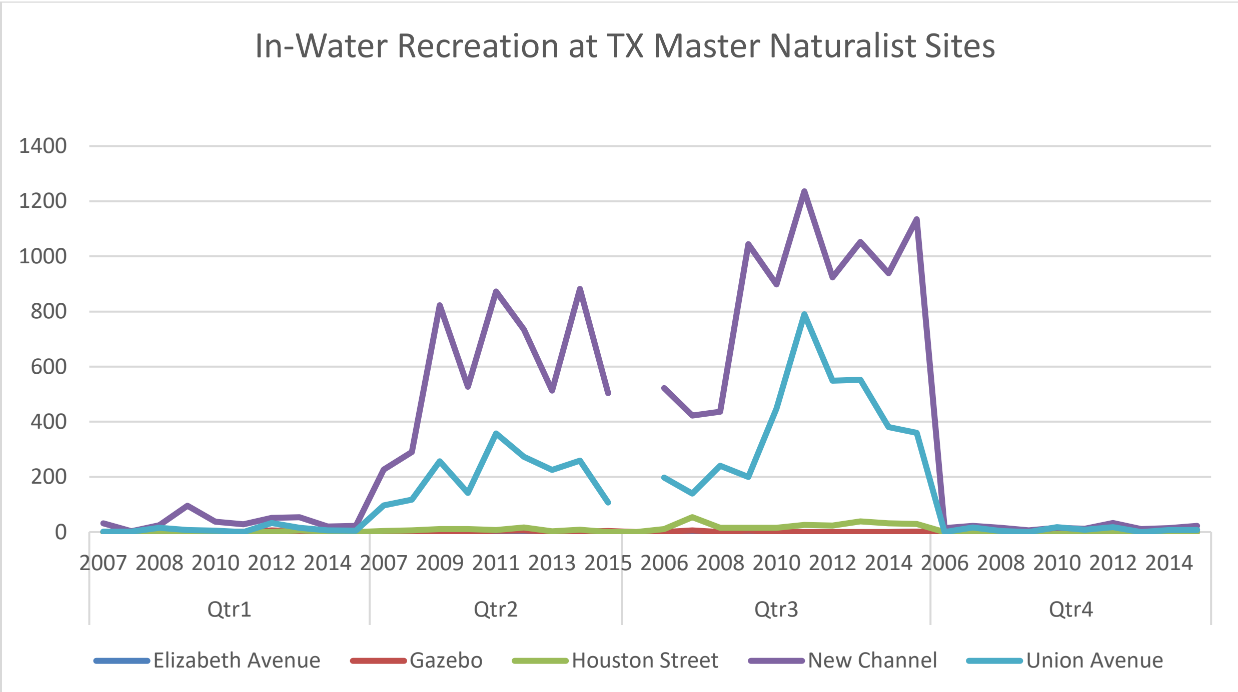

Figure 3.19. Averages of recreational pressure counts per the Texas Master Naturalist dataset by quarter and year.

4 DISCUSSION

4.1 Vegetation

The distinction of sites based on aquatic vegetation was visibly apparent among the two spring systems as shown in Figure 3.1. Though exceptions occurred, there was minimal overlap between sites over time based on the coverage and abundance of plant species.

The most consistent plant community was found at Landa Lake (Figure 3.1). The sampling results over the fourteen-year period yielded a very homogeneous community. This is an expected result due to the unique and stable conditions at the site. Due to the impoundment, this lentic system presents a unique habitat and less physical stress as compared to the lotic sites. This situation allows for distinct plant communities and growth densities. Conversely, the site that exhibited the greatest variance over time was Lower New Channel. This high variance is possibly due to the known recreational pressure.

The evidence for the impact of physical stress due to flow is supported by the analysis show in Figure 3.4 and Figure 3.6 The CCA in Figure 3.4 indicated that three categories of flow measure were found to have significant associations with the abundance and diversity of plant communities between sites at both spring systems. The number of high flow events prior to sampling was the most significant parameter giving indication that high flows create scour strong enough to alter plant communities. A potential stressor to the wild rice due to emergent plant abundance was indicated in the model selection of the NPMR as shown in Figure 3.5. This observed situation would be expected to be related to the emergent plant shading effect on the rice. The significant association of predation, however, exhibited positive association. This trend is not expected to be a cause/effect situation and is probable to be a result of more predation being observed with a greater plant abundance. It has been indicated that these predation survey results are not believed to be entirely valid due to sampling inconsistencies.

4.2 Fish

The existing fish dataset and simultaneously collected environmental variables only allowed for the comparison of the endangered fountain darter with vegetation coverage. However, the results of the analysis indicate that this association was indeed significant. Based on the simple correlation of vegetation coverage percentage and abundance of darters (Figure 3.7), a strong positive correlation appears evident. This result was no surprise as it is well known that fish associate with cover to both avoid predation and to forage for smaller fish. The International Union for the Conservation of Nature (IUCN 2017), Alexander (2012), and Schenck (1976) (among others) specifically document the fountain darter's preference for dense vegetation. The NMS shown in Figure 3.8 indicated preferences of the fountain darter to particular plant species yet that preference varied dependent upon the site. This site-level variance for plant-specific preferences is most-likely a result of the fish using the available plants for cover dependent on their availability at a site. This hypothesis is supported by Figure 3.9 based on the lack of pattern observed between darter numbers and plant species between sites. The most dominant population of darters, however, was found in the site with the most abundant plant diversity and coverage (i.e., Landa Lake). This is consistent with the known preference of the darter to backwater areas (e.g. lentic water) (IUCN 2017, Alexander 2012). Other than habitat and/or prey abundance, this artifact could also be related to the stable conditions of the lake discussed in Section 4.1.

4.3 Macroinvertebrates

Discerning a pattern of variation of the macroinvertebrate communities was somewhat limited due to issues within the existing datasets. Spatial and temporal patterns could not be differentiated due to a lack of discernable variance over time and between sites using the three endangered species count data (NMS on "macro drift" dataset). This was most likely an artifact of the large size of the dataset when the samples from the two springs were combined. In fact, after splitting the data into the two springs systems, a significant direct ordination resulted. Using the limited explanatory variables, the significant RDA model from the Comal Springs indicated that temperature played a strong role in describing the difference in macroinvertebrate diversity. The indication from this graph (Figure 3.10) is that the Western Upwelling site had consistently higher temperatures and the genera *Stygobromus* demonstrated an affinity for these conditions. Conversely, the taxa *Heterelmis* and *Stygoparnus* are found in greatest abundance in the lower temperature sites of Spring Run 1 and 3. Riffle beetles (*Heterelmis*) have been shown to be negatively impacted by increased temperatures (Elliott 2008) whereas amphipods are generally known as very tolerant of varying/poor environmental conditions (Arthur 1982, Tsoi 2006)

The slight implication of a velocity preference was also indicated by the significant, though reduced, association of its vector (arrow). Both *Stygoparnus* and *Stygobromus* were found in slightly greater numbers during years in which Spring Run 1 and 3 and Western Upwelling had faster flow (predominantly 2009/2010). Reinforced by the strong correlation between flow and *Heterelmis* capture rate (Figure 3.11), it does appear that the macroinvertebrates do generally prefer increased flow, which could be associated with increased dissolved oxygen. However, with the existing information it is impossible to determine if this is a "preference" by the invertebrates or an artifact of the environment and sampling method. It is possible that increased flow from the springs could increase capture efficiencies as the extra velocity might produce a flushing effect resulting in more individuals being forced into the trap as compared to low flow conditions. Resolving this unknown is a good subject for additional study.

4.4 Salamanders

The sole dataset for salamanders yielded interesting results however many of these must be discounted due to the known ecology of the site. Though significant correlation was found with regard to predator-prey interactions, water quality concerns, and temperature preferences; issues with biological sampling variance and water quality parameter concentrations under differing flows is believed to be the result of strong predictor/response associations. As shown in Figure 3.12, a positive correlation between the salamanders, *Seborgia* (amphipod), and mites indicates a prevalence for the salamander to associate with areas of highest food sources. However, the abundance of all three taxa is instead likely related to the stressors of flow on each. The even stronger model that focused on NO_3 and dissolved oxygen gave implications that elevated nitrates were negatively impacting salamander abundance, especially since nitrates have a known toxicity to most life in general. Though this association cannot be disproven with the existing datasets, it is more probably that low flow alone was responsible for the concentrated nitrate levels and reduced numbers of salamanders.

Increasing temperature also appeared to have negative associations with salamander abundance. As seen in the NPMR model in Figure 3.14 and Figure 3.15, salamander numbers declined precipitously when temperatures increased. This is not a unique phenomenon. A 2014 study with lungless salamanders in the Appalachian Mountains showed that increased temperatures (both ambient air and water) increased salamander metabolism leading to a quicker exhaustion of the individual's energy stores. It was also found that increased temperatures also resulted in smaller adult body sizes and reduced fecundity (Caruso et al. 2014). However, elevated water temperature would also be expected during periods of reduced flow as pooled water warms to ambient temperatures in the absence of abundant cool spring water. The low flow conditions coincidentally are implicated in either lower overall salamander concentrations or a reduction in the efficiency of trapping, which potentially negates the trends found in this analysis.

4.5 Water Quality

The outliers for water quality parameters occurred consistently throughout the Comal Springs sampling sites during the third quarter of 2014. As indicated in the results section, excessive nitrates were one of the primary water quality measures responsible for this site/year/quarter distinction. The summer of 2014 was a severe drought in southeastern Texas so any inputs of nutrients would become more concentrated due to reduced flow (i.e., less dilution). Potential exists for nutrient inputs related to aging infrastructure and also fertilization of adjacent lawns/golf courses.

4.6 Recreation

Figure 3.19 indicates excessive recreational pressures from 2009-2014 during quarter three at two of the areas sampled for this measure: New Channel and Union Avenue. It is our belief that the excessive nitrate concentrations at Comal Springs result from a combination of the drought conditions in 2014 and the elevated nitrate source from a non-point source.

5 CONCLUSIONS

This study has addressed the primary questions of the research effort utilizing existing datasets, including assessments of the extent of water chemistry variation, primary stressors that exist for populations of concern, and the extent of the variation of species over time and between sampling reaches of Comal and San Marcos spring systems. Future studies could facilitate this analysis by improving the consistency of data collection and the collection of other potential stressors.

Physical disturbance, either through excessive flow or potentially recreational disturbance, were indicated as the primary determinant of plant coverage in the spring systems. No measured stressors were implicated in the diversity of plants other than known differences in habitat (i.e., lentic versus lotic water). Fish clearly depended on adequate vegetation coverage to proliferate yet do not appear to prefer any particular species of plant. Future studies should focus on the ideal amount of vegetation coverage for fish while controlling for the variation of other habitat variables.

The effects of elevated temperature on the endangered beetle population could possibly provide implications for management. A potential sampling bias related to flow should be explored in order to ensure consistent population estimates.

Water quality parameters varied randomly across sampling sites and time with the exception of the 2014 drought due to the concentrating effects of low flow (e.g., elevated nitrates). Though the interaction of the two cannot be proven in this analysis, the threat of excessive nitrates is a concern for the salamander population especially during low flow conditions. Beyond threats due to water chemistry, it is recommended that the diet of salamanders be studied to understand the limitations due to prey abundance.

REFERENCES

Alexander, M.L. & C.T. Phillips. 2012. Habitats used by the endangered fountain darter (*Etheostoma fonticola*) in the San Marcos River, Hays County, Texas. The Southwestern Naturalist, 57(4): 449-452.

Arthur, J.W. 1982. Effect of elevated water temperature on macroinvertebrate communities in outdoor experimental channels. Water Research, Vol. 16(10): 1465-1477.

BIO-WEST 2014a. Habitat Conservation Plan Biological Monitoring Program. Comal River Aquatic Ecosystem 2013 Annual Report. Edwards Aquifer Authority. 92 p. plus Appendices.

BIO-WEST 2014b. Habitat Conservation Plan Biological Monitoring Program. San Marcos River Aquatic Ecosystem 2013 Annual Report. Edwards Aquifer Authority. 80 p. plus Appendices.

Caruso, N.M., M.W. Sears, D.C. Adams, & K.R. Lips. 2014. Widespread rapid reductions in body size of adult salamanders in response to climate change. Global Change Biology, 20: 1751-1759. doi: 10.1111/gcb.12550.

EARIP, 2012. Edwards Aquifer Recovery Implementation Program – Habitat Conservation Plan. Edwards Aquifer Authority. 414 p. plus Appendices.

Elliott, J.M. 2008. The ecology of riffle beetles (Coleoptera: Elmidae). Freshwater Reviews 1(2): 189-203.

Hecnar, S.J. 1995. Acute and chronic toxicity of ammonium nitrate fertilizer to amphibians from southern Ontario. Environmental Toxicology and Chemistry, Vol. 14, No. 12: 2131-2137.

International Union for Conservation of Nature. 2017. *Etheostoma fonticola* habitat and ecology. http://www.iucnredlist.org.

Jongman, R.H.G., C.J.F. Ter Braak, & O.F.R. Van Tongeren. 1995. Data analysis in community and landscape ecology. Pp. 91-173. Wageningen: Pudoc.

Legendre, P. & E.D. Gallager. 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129:271-280.

McArdle, B.H. & M.J. Anderson. 2001. Fitting multivariate models to community data: a comment on distance-based redundancy analysis. Ecology., 82(1): 290-297.

McCune, B. & M. J. Mefford. 2016. PC-ORD. Multivariate Analysis of Ecological Data. Version 7.02. MjM Software, Gleneden Beach, Oregon, U.S.A.

McCune, B. 2009. Nonparametric multiplicative regression for habitat modeling. http://www.pcord.com/NPMRintro.pdf.

McCune, B. & M. J. Mefford. 2009. HyperNiche. Nonparametric Multiplicative Habitat Modeling. Version 2.30. MjM Software, Gleneden Beach, Oregon, U.S.A.

McCune, B. 2006. Non-parametric habitat models with automatic interactions. Journal of Vegetation Science 17: 819-830, 2006.

McCune, B., J.B. Grace. 2002. Analysis of Ecological Communities. MjM Software, Gleneden Beach, Oregon, U.S.A.

Ortiz-Santaliestra, M.E. & M.J. Ferandez-Beneitez, A. Marco. 2012. Density effects on ammonium nitrate toxicity on amphibians. Survival, growth and cannibalism. Aquatic Toxicology, 110-111: 10-176.

Schenck, J.R., & B.G. Whiteside. 1976. Distribution, habitat preference and population size estimate of *Etheostoma fonticola*. Copeia 1976:697-703.

Ter Braak, C.J.F. & P. Šmilauer. 2012. Canoco reference manual and user's guide: software for ordination, version 5.0. Microcomputer Power, Ithaca, USA, 496 pp.

Ter Braak, C.J.F. 1994. Canonical community ordination. Part I: Basic theory and linear methods. Ecoscience, 1(2): 127-140.

Tsoi, K.H. K.M. Chiu, & K.H. Chu. 2006. Effects of temperature and salinity on survival and growth of the amphipod *Hyale crassicornis* (Gammaridea, Hyalidae).