Updated Riparian Functional Assessment, 2014
SR-15-10, April 2015

Aaron Richter, EIT
Ana González, PhD

City of Austin
Watershed Protection Department
Environmental Resource Management Division

Abstract

The City of Austin Grow Zone Program has established a set of degraded sites along urban streams where management is aimed at restoring the riparian vegetation. Monitoring the improvement in ecological function in these degraded urban sites is important to assess the functional improvement of these sites relative to reference urban sites with relatively undisturbed riparian forest. In 2012, a Riparian Functional Assessment tool was developed to address this need and has included a set of modifications based on early results. A set 15 of functional parameters were measured in both degraded and reference sites. Hierarchical model analysis showed that seven of these functional parameters were significantly different in reference relative to degraded sites. Random Forest Analysis was used to examine the importance of each of the functional parameters in correctly classifying a site as 'reference or 'degraded'. In addition, the associated probability of being classified as belonging to a class given the values for the functional parameters is used as a quantitative measure of overall ecological function. Six parameters were found to be important for classifying a site. The RFA index clearly segregates reference and degraded sites. Over time, as degraded sites recover more of their ecological function, the expectation is that their RFA index will be more similar to those of reference sites. This report includes recommendations to incorporate or modify parameters to further improve this functional monitoring tool for riparian areas.

Introduction

A riparian zone is a complex assemblage of plants and other organisms in an environment adjacent to water (Lowrance et al. 1985). As a transition zone between the upland and aquatic ecosystem, riparian zones are diverse communities that possess physical attributes, biotic properties, and energy flow processes unique to the interaction of the surrounding ecosystems (Naiman and Decamps 1997). Because of this unique interaction, riparian zones perform a wide range of ecological functions which affect hydrologic dynamics, water quantity, and water quality of the adjacent aquatic ecosystems (Correll 1999, Groffman et al. 2003). Riparian vegetation can reduce erosion by stabilizing bank material or reducing the velocity of water in the channel (Naiman and Decamps 1997). Woody debris produced from fallen or dead vegetation adds habitat for fish and macroinvertebrates (Anderson and Sedell 1979; Harmon et al. 1986). Riparian canopy provides temperature buffering, which protects the biota from the adverse effects of large fluctuations in temperature (Stacey et al. 2006). Live woody vegetation sequesters carbon while decaying organic material acts as a source of nourishment for aquatic biota (Fischer and Fischenich 2000; Stacey et al. 2006; Richardson et al. 2007; Woolsey et al. 2007). Soils and vegetation filter nutrients from nonpoint sources of pollution (Lowrance et al. 1983; Peterjohn and Correll 1984; Jacobs and Gilliam 1985). The level of function provided depends on the integrity of the riparian zone. In general, the more degraded an ecosystem, the more fundamentally altered the basic ecosystem services it can provide (Hobbs and Cramer 2008). Thus, many comprehensive monitoring programs include riparian zone condition as part of their environmental assessments (Wissmar and Beschta 1998; Barbour et al. 1999) and riparian zone restoration is commonly applied to improve the ecological function at degraded sites.

A vast majority of current restoration endeavors involve either removal of vegetation, planting, or both, without measuring essential ecosystem processes that may be affected. Previous work by the City of Austin (COA) identified specific metrics that should respond to changes in management of the riparian zone as well as represent a broad suite of ecological function (Duncan 2012). The metrics were used in the development of the Riparian Functional Assessment (RFA) tool (Richter and Duncan 2012) which was to be used to quantitatively measure how restoration projects strengthen the environmental functionality of the riparian zone, to provide information for adaptive management of degraded sites, and to track the progression of the vegetation in degraded conditions towards an achievable reference condition.

Soil compaction, soil moisture, in-stream cover, riparian zone width, plant cover and structural diversity, hardwood demography, and recruitment were the parameters used in the initial RFA tool. Increased compaction can reduce the soil's ability to function for structural support, water and solute movement, and restrict root growth (USDA 2008). This leads to shallow rooted plants and poor plant growth, reduced vegetative cover, increased erosion, and reduction in water infiltration (USDA 2008). Soil moisture was collected to try and capture changes to hydrology associated with urbanization which can result in lower water tables and drier more aerobic soil conditions (Gift et al. 2010). These changes can result in reduced denitrification and altered plant species composition (Gift et al. 2010, Sung et al. 2011). In-stream cover was collected as a surrogate to diel water temperature monitoring. The amount of solar shading provided by adjacent and in-stream riparian vegetation is critical for maintaining temperature buffering. A lack of cover leads to increased solar exposure of the channel and can cause changes in water temperature over a diel period that negatively impacts the aquatic life of the stream (Stacey et al 2006). Riparian zone width was collected to represent the buffer width for a stream as the riparian zone has been shown to filter pollutants, control erosion, prevent flooding, and provide habitat and nutrient inputs into the stream (Barbour et al. 1999, Fischer and Fischenich 2000). High amounts of plant cover and structural diversity of vegetation (groundcover, understory, and canopy) are indicative of a productive plant community, high species diversity, adequate food resources and habitat for wildlife, and reduced flood impacts along banks (Stacey et al. 2006). Demography of the dominant hardwood species was noted to indicate recruitment of the dominant tree species and disturbance intervals of a site. The dominant hardwood species was tracked because the dominant tree species exerts the most influence on the function of an area, thus the greatest functional changes will occur if the abundance of the dominant species is altered (Richardson et al. 2007). Special emphasis was given to recruitment of the dominant species because the sapling community reflects current ecological condition of habitat while mature communities are reminders of past environmental condition (Woosley et al. 2005). Follow up analysis of data collected after one year of riparian restoration raised concerns about the suite of environmental variables represented in the RFA tool (González and Richter 2014).

Soil moisture was replaced by the soil organic carbon content, demography and recruitment were altered to incorporate all species, and the following environmental metrics were added to data collection: count of large woody debris, count of snags, and presence of hydrophytic vegetation. Instantaneous soil moisture measurements were found to be influenced by the time of day that sampling occurred and were not informative. Organic carbon present in the soil will retain water and thus the soil organic carbon and the long-term soil moisture are positively correlated (Rawls et al. 2003). Thus, soil organic carbon was collected as a surrogate to soil moisture.

Hardwood demography of the dominant species does account for the most influential species at a site but would potentially treat certain natural processes, such as forest succession, as a negative impact to a site. This is an undesirable result of the RFA tool. Therefore, the presence of every tree species was collected in three size classes: seedling, sapling, and mature. This allowed for the analysis of not only the size class distribution of species to determine the self-sustainability and disturbance level of a community but also allowed for an analysis of the species richness and wetland status. Species richness, the number of species present, is a simple way to describe community diversity (Magurran 1988). High species diversity has been found in sites that provide essential ecosystem services while research suggests ecosystem function declines when species diversity is reduced to low levels (Moffatt et al. 2004, Richardson et al. 2007). The wetland status of species is categorized as OBL (almost always a hydrophyte), FACW (usually a hydrophyte but occasionally in uplands), FAC (commonly occurs as either a hydrophyte or non-hydrophyte), FACU (occasionally a hydrophyte but usually in uplands), and UPL (rarely a hydrophyte, almost always in uplands). Species were assigned a wetland status by the US Fish and Wildlife in 1988, but the US Army Corps of Engineers has since provided an updated plant list (Lichvar and Kartesz 2009). Sites with a majority of OBL species should be saturated more consistently than sites that consist mainly of UPL species and should provide more essential ecological function. The presence of hydrophtic vegetation was also monitored as hydrophytic plants are indicators of high soil moisture and are sensitive to groundwater decline in semiarid regions (Stromberg et al. 1996).

Snags are difficult to correctly identify to species in trees that have been dead for a long time and no longer contribute to the demographic dynamics of the species present. Therefore, snags were removed from the demography classification and added as a separate metric where the number of snags present were counted but not identified. Snags provide important ecological services to riparian forests (Groffman et al. 2003, Gurnell et al. 2005) and represent a potential source of large woody debris for the stream. The functional contributions from large woody debris are well documented and include greater habitat diversity for aquatic life, more natural channel geometry, and greater resistance to flood events (Abbe and Montgomery 1996, Hyatt and Naiman 2001, Larson et al. 2001, Stacey et al. 2006). Large woody debris was considered in the original RFA tool metric list but was not a chosen metric due to its variation in undisturbed sites (Richter and Duncan 2012). It is possible that the differing drainage area at reference sites was influential on the large woody debris count and has been accounted for in the current sampling scheme.

One objective of the RFA tool is to track the above metrics at disturbed sites and determine how the metrics change as restoration progresses. Each metric mentioned can assist in the description of a disturbed site and what ecological function may be lacking at a site. However, measuring the success of restoration based on all of the above metrics is a faulty method as some of the metrics may not change through the restoration process. A portion of the development of the RFA tool includes choosing desirable metrics to track that should change over time as a site undergoes restoration.

Goals of restoration should include a desired status of a site with a way to measure the success of a restoration effort towards that desired status. The desired status for a site should be achievable and within the constraints of the project location. Disturbed sites tracked by the RFA tool are urban riparian zones. Urban forests are altered from pristine conditions regardless of the level of continuous disturbance present. The achievable level of ecological function that can be restored to the disturbed sites is a level equal to the ecological function present in undisturbed urban riparian areas. Thus, metrics for the RFA tool are also collected within undisturbed urban riparian areas as a reference towards which progress during restoration is compared.

Desirable metrics to track the success of restoration should be those metrics that are important in defining the distinction between the undisturbed 'reference' riparian areas and the disturbed 'degraded' sites. Metrics for which there is no distinguishable difference between the two groups of sites are probably not metrics that can be used to assess the trajectory of a site towards a reference condition either because management practices cannot alter them significantly or because their inherent variation is such that it prevents detecting significant changes without unrealistic sampling effort. A classification model can be constructed using the reference riparian sites as one group and the degraded riparian sites as a second group. Metrics which lead to an accurate classification model would be considered important metrics in the distinction between the reference sites and the degraded sites. These metrics could then be used in the RFA tool to track a degraded sites progress towards an achievable level of function.

Random forest analysis is a type of classification tree within the realm of statistical learning which can identify important predictor variables in regards to a response variable (Breiman 2001). Measured responses can be quantitative or qualitative. Classification trees are powerful tools which partition a data space into small groups and then make simple predictions within the partitioned groups. Data is partitioned at nodes based on the values of a predictor variable. Each data point travels down a set of nodes going either left or right based on the value of the predictor variable at that node until a leaf node is reached at which point a prediction of the response variable is made. Random forest analysis forms multiple classification trees on the data set and has been shown to have lower misclassification rates when compared to single classification tree techniques (Breiman 2001).

More specifically the random forest algorithm is as follows:

- A bootstrapped sample of the data set is drawn with replacement from the data. About one third of the data is left out to form the OOB (out-of-bag) data for a particular tree.
- A random forest tree is grown using the bootstrapped data. A tree is grown by selecting a random set of predictor variables from the total list of predictor variables, picking the best variable from the randomly selected variables to split the data, and splitting the data to two 'daughter' nodes. This process continues until a minimum node size is reached.
- The above steps are repeated to form separate bootstrapped data sets and random forest trees.
- The categorical response predicted from a random forest analysis for each data point is the predicted categorical response in the majority of random forest trees.

The decision to split data into two 'daughter' nodes is done using the Gini criterion (Breiman 2001, 2002). The function for the Gini criteria for classification is:

$$gini(t) = 1 - \sum_{j=1}^{K} p\left(c_j \middle| t\right)^2$$

where t is a node, c_j is a classification, and K is the total number of classifications. The probability of a data point being in class j given the outcome of a test at node t is then $p(c_j|t)$. The selection process in random forest analysis attempts to select the variables at a given node that maximize the above function, which should output two 'daughter' nodes with the most accurately classified data points in each. The Gini criteria is one of the most popular split functions in decision tree analysis. The other most popular function appears to be Information Gain function taken from Information Theory and entropy of a system. Little empirical evidence has been found that these two split functions produce differing classifications and theoretical work has shown that in a sequence of tests the disagreement in classification between the two functions was never higher than 2% disagreement (Raileanu and Stoffel 2004).

The out-of-bag data sets computed during the random forest process are important for error estimates and variable importance estimates. Typically in a classification model, a certain amount of data is set aside for cross-validation of the classification model. During this process, a model produces predictions for classification and those predictions are compared to known classifications for each data point. A misclassification rate can be calculated during this cross-validation process. In random forest analysis, each model formed from bootstrapped data predicts the classification for the OOB data points and calculates an error rate which is the OOB error estimate. This eliminates the need to set aside data points for cross-validation in a random forest analysis (Breiman 2001).

In random forest analysis there are four measures used to calculate a predictor variable's importance to the classification (Breiman 2002). For measure one, OOB data is used to randomly permute values of each variable while other variables are held constant. The importance of each variable is computed by comparing how much the OOB error estimate increases for each variable when compared to the original test set error. Measure two is calculated as the average lowering of the 'margin' when a variable is permuted as in measure one. The 'margin' is defined as the proportion of votes for a data point's true class minus the maximum of the proportion of votes for each of the other classes. Measure three is the count of the number of 'margins' that are lowered minus the count of 'margins' raised. Finally, measure four is calculated as the sum of all decreases in the Gini criteria, which is used to split the data, for a given variable normalized by the number of trees. If a variable can be easily classified as an important variable to the classification model then measure one can easily identify the variable as important but, if the number of important variables to the model grows large in a small data set, it may be necessary to calculate and examine all four measures (Breiman 2002).

The current analysis uses a categorical classification of a 'degraded' riparian zone to denote the disturbed sites and a 'reference' riparian zone to denote the undisturbed sites. Random forest analysis was used to build an accurate classification model using the collected environmental metrics as predictor variables in the model. The importance of each metric to the classification model was calculated using measure one of variable importance mentioned above in which the importance is computed as the change in the OOB error estimate compared to the original test set error. Metrics which had a high measure of importance were then used in the calculation of an updated RFA tool.

The RFA tool is a multi-metric index developed and updated to assess the ecological function of disturbed riparian zones managed to restore the riparian buffer. Multi-metric indices in ecology combine information from independent measurements into a single index which is often tied to human-caused disturbances of the environment (Karr and Chu 1997). Multi-metric indices have been shown to be useful as they allow for a quantitative measure of disturbance in the presence of uncertainty of the causal processes and can be used as a starting point for the building of causal models (Riseng et al. 2006). However, there is some criticism to using multi-metric indices as a measure of environmental health because the indices can lack interpretability or usefulness (Suter 1993, von der Ohe et al. 2007). The RFA was designed to measure the environmental functionality of a riparian zone by measuring how similar a set of metrics collected at a degraded site was to the same set of metrics collected within a group of reference sites. More specifically, the probability of a degraded site to belong to a group of reference sites conditional on the metrics collected. We believe this to be a clear and useful way to compress multiple metrics into a single index.

Bayesian decision theory uses Bayes rule to calculate the probability of an object belonging to a particular class. In a two class system, an object would belong to class one if the probability of the object belonging to class one was higher than the probability of the object belonging to class two (Alpaydin 2010). While the general classification of a site is of little interest due to the fact that we have already classified sites in the random forest analysis, the probability related to an object belonging to a particular class is much more interesting. The probability of an object belonging to a given classification is conditional on a set of variables describing that object. A higher probability of belonging to a given classification exists if the values of the conditional variables are similar to typical values within a classification. Thus, a degraded site will have a higher probability of being classified as a reference site if values of the environmental metrics present at a site are similar to typical values of those environmental metrics seen within the reference sites. In addition, as the values of environmental metrics of a degraded site approach values of reference sites over time then the probability of that site to belong to the reference classification should increase over time.

In general, the probability of a site to belong to a classification *c* given a set of environmental metrics *x* can be denoted as $P(c|\boldsymbol{x})$. Bayes rule is then used to solve for this probability:

$$P(c|\boldsymbol{x}) = \frac{P(\boldsymbol{x}|c)P(c)}{P(\boldsymbol{x})}$$

In the above formula $P(c)$ is the prior probability of a site to be in a specific classification, $P(\boldsymbol{x}|c)$ is the likelihood of observing the vector $\boldsymbol{x}$ given that the site is in a specific classification, and $P(\boldsymbol{x})$ is the probability of observing the vector $\boldsymbol{x}$. More specifically, the classification for the RFA consists of two classifications and the probability of observing the vector $\boldsymbol{x}$ can be simplified into:

$$P(\boldsymbol{x}) = \sum_{i=1}^{2} P(\boldsymbol{x}, c_i) = \sum_{i=1}^{2} P(\boldsymbol{x}|c_i)P(c_i)$$

which is the probability of vector $\boldsymbol{x}$ being observed given the site is in classification one multiplied by the probability that the site is in classification one plus the probability of vector $\boldsymbol{x}$ being observed given the site is in classification two multiplied by the probability that the site is in classification two.

$P(\boldsymbol{x}|c_i)$ can be more difficult to calculate. However, using the naïve Bayes classifier, if $\boldsymbol{x}$ is a vector of *N* independent environmental metrics then the likelihood of observing the vector $\boldsymbol{x}$ given that the site is in the *i*th classification can be written as:

$$P(\boldsymbol{x}|c_i) = \prod_{j=1}^{N} P(x_j|c_i)$$

which is the product of the probability of observing the value of the individual environmental metrics given the classification state. There were no strong correlations between environmental metrics collected in the RFA for 2014 when the data was divided between degraded and reference sites. Thus the environmental metrics will be considered locally independent and naïve Bayes classification will be used to calculate the likelihood of observing the vector $\boldsymbol{x}$ of environmental metrics given the classification is in either degraded or reference.

The posterior probability of being classified as a degraded site or a reference site will then be calculated for each site given the values associated with each environmental metric.

$$\mathrm{P}(\mathrm{c_{ref}}|\mathrm{x}) = \frac{\mathrm{P}(\mathrm{c_{ref}}) \prod_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{P}\left(\mathrm{x_j}\middle|\mathrm{c_{ref}}\right)}{\mathrm{P}(\mathrm{c_{ref}}) \prod_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{P}\left(\mathrm{x_j}\middle|\mathrm{c_{ref}}\right) + \mathrm{P}(\mathrm{c_{deg}}) \prod_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{P}(\mathrm{x_j}|\mathrm{c_{deg}})}$$

$$P(c_{deg}|x) = \frac{P(c_{deg}) \prod_{j=1}^{N} P(x_j|c_{deg})}{P(c_{ref}) \prod_{j=1}^{N} P(x_j|c_{ref}) + P(c_{deg}) \prod_{j=1}^{N} P(x_j|c_{deg})}$$

where N is the number of environmental metrics used in the calculation. For the initial calculation of the posterior probabilities the prior probabilities will be uninformative, so P(cref) = P(cdeg) = 0.5. P(cref|x) is the probability of a site to belong to the group of reference sites given the collected environmental metrics at the site and shall be the basis of the RFA index.

Methods

Study Area

The study area is located within Austin, TX, and sampling sites constrained within the boundaries of 13 km west of Mopac Expressway, 13 km east of IH-35 and the north and south boundaries of Travis County (Figure 1). Two ecoregions, the Edwards Plateau and the Blackland Prairie, are defined within the study area. Land use is mostly urban in the center of the area and more suburban towards the edges with urban-rural interfaces. Average maximum temperature in August is 36.1°C and average minimum temperature in January is 5.3°C. Mean annual precipitation is 870 mm with high temporal variation, although May and June are generally the wettest months and July is the driest. Averages are calculated from Local Climatological Data (LCD) collected by the National Weather Service at the Austin City/Camp Mabry site from 1981-2010. Upland vegetation in both ecoregions is better described as a dynamic mosaic of woodlands, savannas with varying degrees of woody cover, and grasslands mostly devoid of woody plants (Fowler and Simmons 2008). Riparian vegetation in the Edwards Plateau ecoregion is often a fully developed forest characterized by *Acer negundo*, *Diospyros texana*, *Garrya ovate* ssp *lindheimeri*, *Ilex vomitoria*, *Juniperus ashei*, *Celtis laevigata*, *Fraxinus sp*, and *Cornus drummondii.* Fully developed riparian forests are less common in the Blackland Prairie portion of the study area given that its land use history has been dominated by agriculture. Where present, riparian forests in this ecoregion are characterized by *Forestiera pubescens*, *Ilex decidua*, *Sapindus saponaria* var *drummondii* , *Celtis laevigata*, *Fraxinus sp*, *Cornus drummondii*, and *Ulmus crassifolia* (Duncan et al. 2011).

Sampling Sites

Twelve sites with a relatively long history of undisturbed vegetation were selected as reference sites (Figure 1), three for each of the following drainage area categories: 0-64 acres, 64 – 320 acres, 321-640 acres, and 641 to 1280 acres (0-26 ha, 26.1 to 129.9 ha, 130 to -259 ha, and 259.1-518 ha, respectively). These categories were defined to coincide with drainage area breakpoints for COA watershed regulations. Twenty four sites under passive or active riparian restoration were selected as degraded sites; half sampled in even numbered years and half in odd numbered years. Most reference and degraded sites fall within a transitional area between the Edwards Plateau and the Blackland Prairie ecoregions. In April of 2014, a set of functional parameters were collected in all reference sites and ten of the twelve degraded sites scheduled for sampling on even years (two sites were undergoing construction). In 2015, data will be collected for all reference sites and a second set of ten degraded sites.

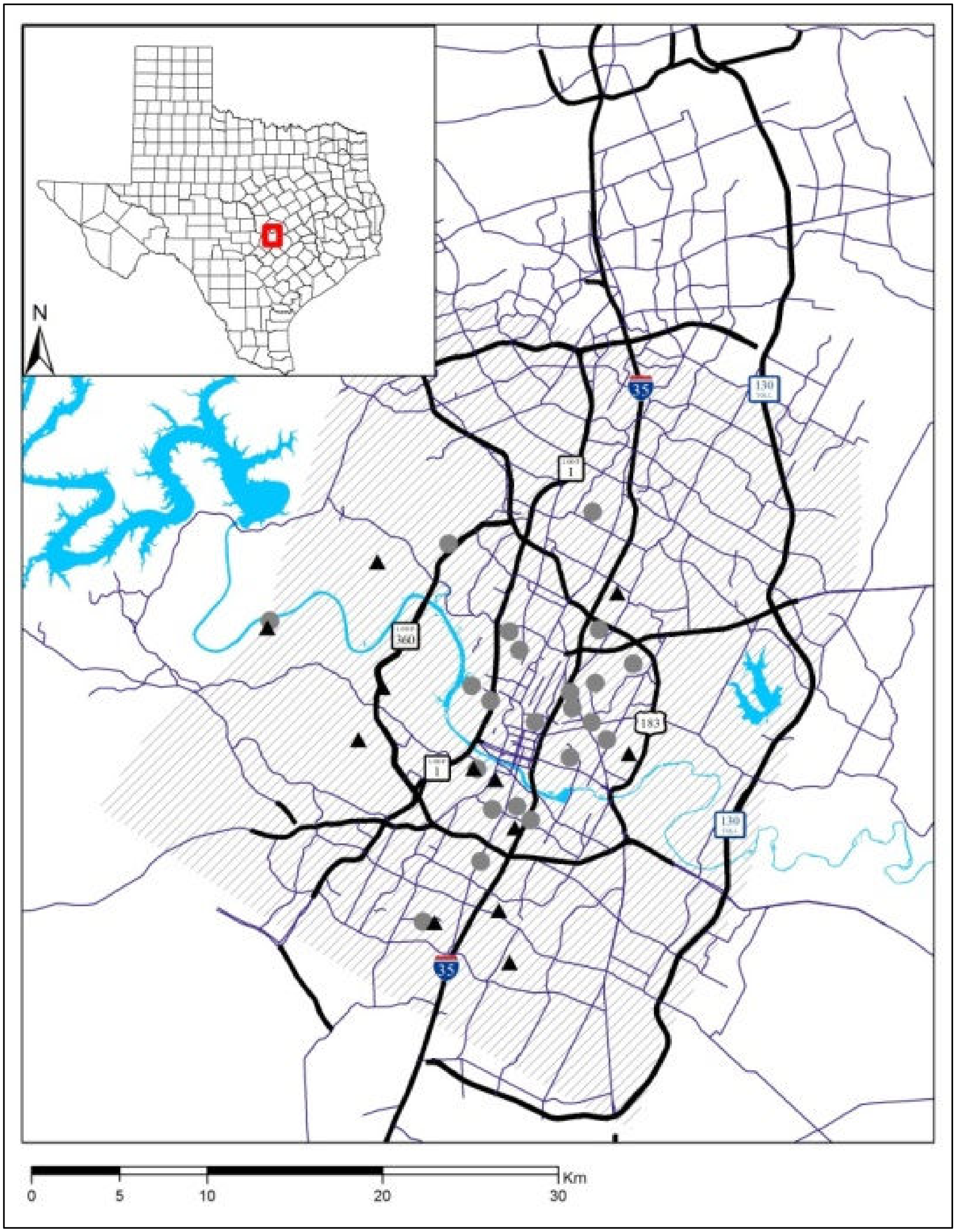

Figure 1: Study sites where triangles represent reference sites and circles represent degraded sites.

Table 1: Riparian Functional Assessment sampling sites. Rows in gray represent degraded sites.

Site Number	Site Name	Drainage Range (acres)	Sampling year
5301	Seabrook Spring	0-64	even
5810	Harper's Branch at Parker Park	0-64	even
5601	Walnut Trib at North Star Greenbelt	0-64	even
5598	Taylor Slough South in Reed Park at Footbridge	65-320	even
5556	East Bouldin at gabion in Gillis Park	65-320	even
5584	Buttermilk at Buttermilk Park	65-320	even
5595	Tannehill at Bartholomew Park near Berkman Dr	321-640	even
5811	Shoal at Crestmont Park	321-640	even
10447	Waller Creek ds 55th Half	321-640	even
5588	Common Ford ds crossing in Common Ford Ranch	641-1280	even
5594	Shoal at Shady Oak Court	641-1280	even
5596	Tannehill us storm pipe in Givens Park	641-1280	even
5580	Barton Trib at Lund and Robert E. Lee	0-64	odd
10877	Eanes at Zilker Disc Golf	0-64	odd
5649	Oak Springs Trib ds Tillery St	0-64	odd
5354	Boggy at Airport	65-320	odd
4835	Boggy at Willowbrook Huisache Crossing	65-320	odd
5591	Johnson in Tarrytown Park	65-320	odd
5582	Blunn at Rosedale	321-640	odd
5593	South Boggy at Dittmar Park near Strickland	321-640	odd
3537	Little Walnut us Mearns Meadow	321-640	odd
5585	Boggy at 10th St	641-1280	odd
3794	Bull at Bull District Park (Lakewood Dr)	641-1280	odd
5592	Little Walnut at Dottie Jordan Park	641-1280	odd
633	Barton Creek Ephemeral 3	0-64	all
5605	Wililamson at Wagon Bend Trl	0-64	all
10413	Barton at key west from Cape Coral	0-64	all
5581	Bee ds Loop 360	65-320	all
5583	Blunn us Cow Trough Spring	65-320	all
10415	Onion at Nuckols and Thaxton	65-320	all
5599	Little Walnut Trib at Gus Garcia Park	321-640	all
10412	West Bull at Long Canyon and Standing Rock	321-640	all
10414	South Boggy at Latteridge Almondsbury	321-640	all
5589	Commons Ford us bridge at Commons Ford Ranch	641-1280	all
5604	West Bouldin in West Bouldin Greenbelt	641-1280	all
5687	Fort Branch at Tura Ln LISI 1	641-1280	all

Data collection

In each study site, a 100 m transect was defined parallel to the stream centerline. Three 10x10 m plots were marked on each bank at 0-10, 45-55, and 90-100 m along the transect, beginning at bankfull and perpendicular to the stream (Figure 2).

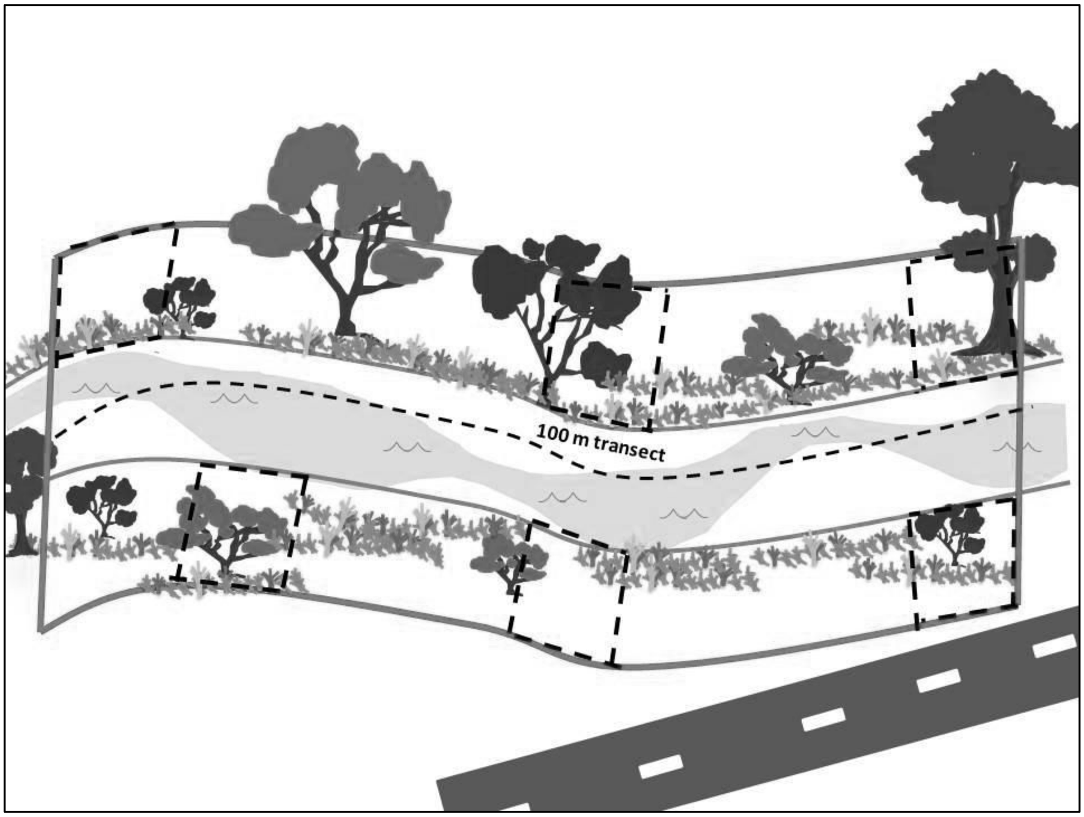

Figure 2: Data collection layout within a site.

Soil Compaction

At the center of each plot, soil compaction was measured with a Humboldt Corps of Engineers Cone Penetrometer at three different points and each measurement was recorded.

Plant Structural Distance

Within each plot, the percent plant cover was visually estimated and recorded for each vegetation layer: canopy (vegetation at ≥5 m), understory (0.5 m<x<5 m), and ground cover (<0.5 m). A vector of the form [x_1, x_2, x_3] was created for each plot where canopy represented x_1, understory represented x_2, and ground cover represented x_3. A weighted Euclidean distance was calculated for each plot from the zero vector [0, 0, 0], using the following equation:

$$Structural\ Distance = \sqrt{3{x_1}^2 + 2{x_2}^2 + {x_3}^2}$$

Weights were assigned to the canopy and understory layers as they are thought to provide higher functionality then the ground cover and represent later stages in a forest succession.

Total Woody Species Richness, Woody Species Richness Distance, and Forest Wetland Affinity
All species of trees present within each demographic category; mature, sapling, or seedling, were recorded. Total woody species richness was calculated for each plot and a species richness vector of the form [y_1, y_2, y_3] was created for each plot where mature trees represented y_1, saplings represented y_2, and seedlings represented y_3. The Euclidean distance was calculated for each plot from the zero vector [0, 0, 0], using the following equation:

$$Richness\ Distance = \sqrt{{y_1}^2 + {y_2}^2 + {y_3}^2}$$

The wetland status indicator of each of the woody species was recorded from the USDA plant database. Obligates were assigned a value of five, FACW species a four, FAC species a three, FACU species a two, and upland species were assigned a value of one. The Forest Wetland Affinity was calculated as the average wetland status indicator and was calculated for each plot.

Riparian Zone Width
On both banks at 5, 50, and 50 m on the transect, the width of the riparian zone, defined in this study as the zone with vegetation left undisturbed for at least a full growing season, was measured beginning at bankfull and perpendicular to the stream centerline.

In-stream Cover
At the center of the channel, canopy cover projected over the stream was measured at 5, 50, and 95 meters with a concave densiometer facing downstream.

Snags
All snags, dead standing trees with a minimum of 20 cm in diameter, present within 10 m area from bankfull along the 100 m transect, were counted.

Large Woody Debris
All pieces of trees with diameter ≥ 15 cm and length ≥ 1 m present within the channel were counted.

Organic Soil Carbon
A sample of approximately 100 g of soil was excavated from the center of each plot after removing the recently deposited leaves and plant debris. All soil samples from all plots in one site were combined into a single composite sample for the site, stored in a one gallon re-sealable plastic bag, and kept in ice. Samples were submitted to the Texas A&M Soil Lab and tested for organic soil carbon through a standard combustion method.

Hydrophytic Vegetation
At bankfull depth, the presence of hydrophytic vegetation, defined as species with $FACW^-$, FACW, $FACW^+$ or OBL wetland status, was noted at 5, 50 and 95 m along the transect.

Data Analysis
Bayesian hierarchical models were used to determine the effect of disturbance on soil compaction, structural distance, total species richness, richness distance, average wetland status indicator, riparian zone width, in-stream cover, number of snags, number of pieces of large

woody debris, and soil organic carbon. Each model used a normal distribution for the response variable except when examining the number of snags and pieces of large woody debris. A Poisson distribution was used to represent the random variation in snags and large woody debris as both are non-negative and discrete variables. Initial models were constructed with fixed intercepts and slopes. Random-coefficients and slopes for each site or drainage area were built into further models. The model with the lowest DIC value for each environmental metric is represented in this report. A site-occupancy model was constructed to determine the effect of disturbance on the presence of hydrophytic vegetation. The above analysis was completed in OpenBugs through R using R2OpenBugs.

For each site the soil organic carbon, average soil compaction, average structural distance, average species richness, average richness distance, average wetland status indicator, average riparian zone width, average in-stream cover, number of snags, pieces of large woody debris, and presence of hydrophytic vegetation were used as input for a random forest model. Analysis was run in R using the randomForest procedure. The number of metrics used at a node in the classification trees was set to one and the number of trees generated was set to 5000. OOB data was used to randomly permute values of each environmental metric while other metrics were held constant. The importance of each metric was computed by comparing how much the OOB error estimate increased when compared to the original test set error. This is denoted as the mean decrease in accuracy of each metric and was plotted.

During the initial computation of the RFA index, the $P(c_{ref})$ and $P(c_{deg})$ are assumed to be equal to 0.5 and the equation for the probability of belonging to the reference class given the values of each metric becomes:

$$P(c_{ref}|x) = \frac{\prod_{j=1}^{N} P(x_j|c_{ref})}{\prod_{j=1}^{N} P(x_j|c_{ref}) + \prod_{j=1}^{N} P(x_j|c_{deg})}$$

The first step to solve this equation is to find the P(xj|cref) and P(xj|cdeg) for every environmental metric. This was done using the density procedure in R which computes kernel density estimates, a non-parametric estimation of the probability density function (probability mass function for discrete metrics) of a random variable, for each metric in reference sites and degraded sites separately. The bandwidth method used followed the methods used by Sheather and Jones (1991). The Gaussian smoothing kernel was used for all continuous metrics while the triangular smoothing kernel was used for discrete metrics. The probability density function or probability mass function was plotted for each metric that was important following random forest analysis. The $P(x_j|c_{ref})$ and $P(x_j|c_{deg})$ for each metric at each site were then drawn from the separate kernel density estimates. To keep these numbers from being too close to zero, each $P(x_j|c_{ref})$ and $P(x_j|c_{deg})$ was log transformed and summed to give the following formula:

$$\log\left(P(c_{ref}|x)\right) = \frac{\sum_{j=1}^{N} \log\left(P(x_j|c_{ref})\right)}{\sum_{j=1}^{N} \log\left(P(x_j|c_{ref})\right) + \sum_{j=1}^{N} \log\left(P(x_j|c_{deg})\right)}$$

The RFA index would then be equal to the $P(c_{ref}|x)$. An assumption of the naïve Bayes classifier is that the space that the two class distributions occupy is stationary. However, it is suspected that as the degraded sites undergo restoration, the metric distributions of the degraded class will shift towards the distribution of the reference class. This would shrink the max value of $P(c_{ref}|x)$ from one when all distributions of each metric are disjoint between the degraded and reference class to 0.5 if all of the degraded and reference distributions overlapped perfectly. Each site $P(c_{ref}|x)$ was multiplied by a correction factor that ranged from one to two was used to correct for this shift in distribution. The Hellinger Distance is commonly used to measure distributional divergence (Basu et al. 1997, Kailath 1967), or the distance and shape difference between two distributions:

$$\text{Hellinger Distance}(P, Q) = \sqrt{\frac{1}{2}\int\left(\sqrt{\frac{dP}{dx}} - \sqrt{\frac{dQ}{dx}}\right)^2 dx}$$

where P and Q are two distribution functions. This form of the Hellinger Distance ranges from zero when distributions are similar to one when distributions are disjoint. Other renditions of the equation leave off the one half multiplier outside of the integral and allow the distance to range from zero to the square root of two. The Hellinger Distance between reference and degraded distributions for each metric were combined using the assumption that the metrics were independent. The final RFA index for each site was calculated using the following equation:

$$RFA = 100 * \left(P(c_{ref}|x) * (2 - HD)\right)$$

where HD is the site Hellinger Distance.

Results

<u>Hierarchical Models</u>

Soil Compaction

The estimate for soil compaction in non-disturbed or reference sites was 146.9 psi[1] (10.33 kg/cm^2) (Table 2). The slope in the model estimated the change in soil compaction when going from a reference site to a degraded site. There was high variability of soil compaction within sites (Figure 3) but soil compaction in degraded sites was estimated to be about 51.5 psi (3.62 kg/cm^2) higher than in reference sites. The credence interval for the slope does not overlap zero so this would be considered a significant difference. Model output that includes intercept shifts for each site can be seen in Appendix A.

Table 2: Model estimates, standard deviation (SD), and credence intervals for soil compaction (psi). The slope shows the increase from the reference group to the degraded group.

	Estimate	SD	Credence Interval (95%)	
Intercept	146.9	5.1	136.1	156.5
Slope	51.5	7.7	36.6	66.8

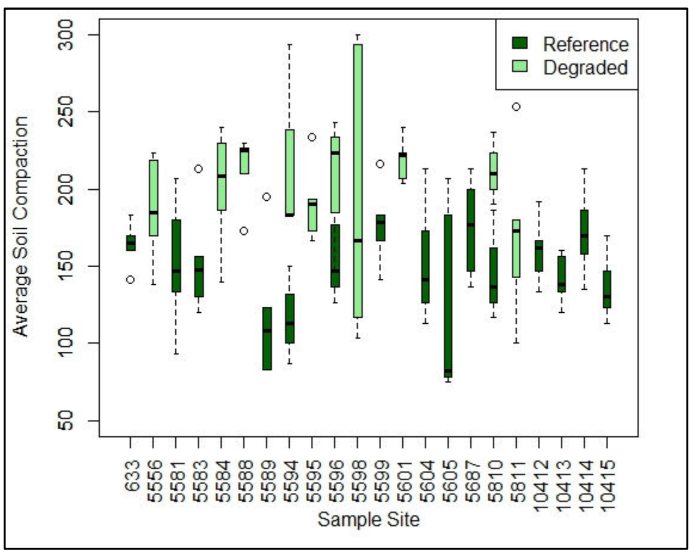

Figure 3: Average plot soil compaction at each site. Sites shown as both reference and degraded correspond to those on which one side of the creek is considered reference and the opposite side as degraded.

Plant Structural Distance

The estimate for the plant structural distance in reference sites was 137.7 (Table 3). The maximum value this number could take would be approximately 245 if there was 100% cover in the canopy, understory, and ground cover. Due to the interaction between the three vegetation layers, light competition from canopy limiting the understory and groundcover layers, this number will never be reached. However, it is enough to know that higher cover in the canopy, understory, and ground cover is represented by larger distance values. Structural distance from zero vector [0,0,0] was estimated to be significantly shorter in degraded sites, indicating that cover in the three layers was less abundant in the degraded class. There was high variability

[1] Instrument measures compaction in pounds per square inch (psi), the SI unit of kg/cm^2 follows in parentheses.

among sites (Figure 4) and it was necessary to allow for site specific adjustments within the model (Appendix B).

Table 3: Model estimates, standard deviation (SD), and credence intervals for plant structural distance. The slope shows the decrease from the reference group to the degraded group.

	Estimate	SD	Credence Interval (95%)	
Intercept	137.7	5.9	125.2	148.5
Slope	-32.9	8.4	-48.4	-15.3

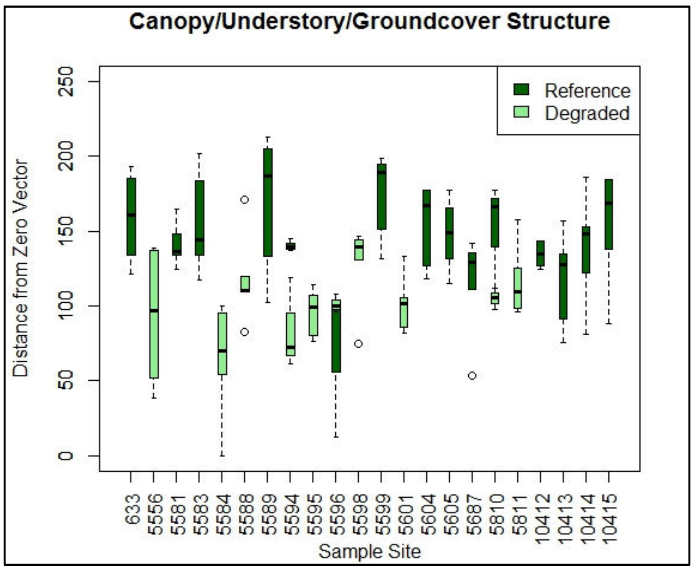

Figure 4: Plant structure [canopy, understory, ground cover] distance from [0, 0, 0] at each site. Sites shown as both reference and degraded correspond to those on which one side of the creek is considered reference and the opposite side as degraded.

Total Woody Species Richness

The estimate for the total woody species richness within the reference group was ten species while the estimate for the total species richness within the degraded group was approximately six species (Table 4). This is the simplest measure of species diversity, thus the diversity was significantly higher at undisturbed sites. Variability was high between sites and within some sites (Figure 5). It was necessary to allow for site specific adjustments within the model, especially site 5583 which had the highest number of species present (Appendix C).

Table 4: Model estimates, standard deviation (SD), and credence intervals for total species richness. The slope shows the decrease from the reference group to the degraded group.

	Estimate	SD	Credence Interval (95%)	
Intercept	10	0.6	9	12
Slope	-4	0.8	-5	-2

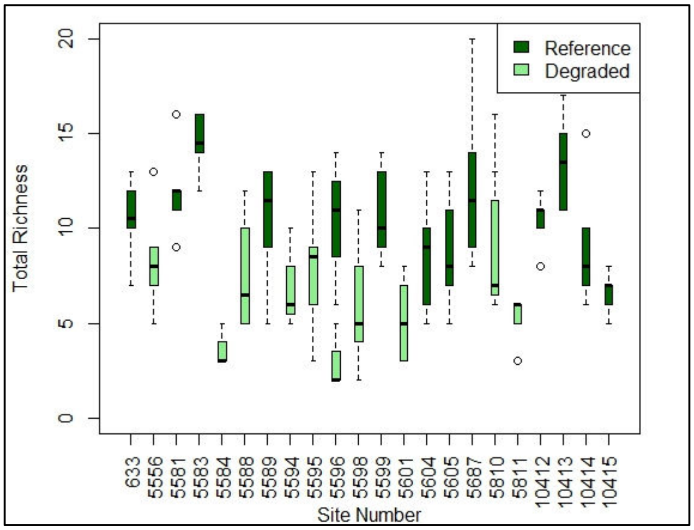

Figure 5: Total species richness at each site. Sites shown as both reference and degraded correspond to those on which one side of the creek is considered reference and the opposite side as degraded.

Woody Species Richness Distance

The estimate for the species richness distance within the reference group was 3.9 and estimated to be 1.1 higher than the distance in the degraded group (Table 5). The distance measures how many species exist within each age class (mature, sapling, seedling). A shorter distance value can equate to one or all age classes containing less species at a site. For this dataset, the shorter distance value equates to a difference in species richness in all age classes as the average species richness in the reference group was higher than the average species richness in the degraded group for the mature trees, saplings, and seedlings (4, 8, 5 vs 2, 5, 2). Variability was high between sites within each group (Figure 6). It was necessary to allow for site specific adjustments within the model, especially site 5583 which had high species richness in all age classes (Appendix D).

Table 5: Model estimates, standard deviation (SD), and credence intervals for species richness distance. The slope shows the decrease from the reference group to the degraded group.

	Estimate	SD	Credence Interval (95%)	
Intercept	3.9	0.1	3.7	4.2
Slope	-1.1	0.2	-1.5	-0.7

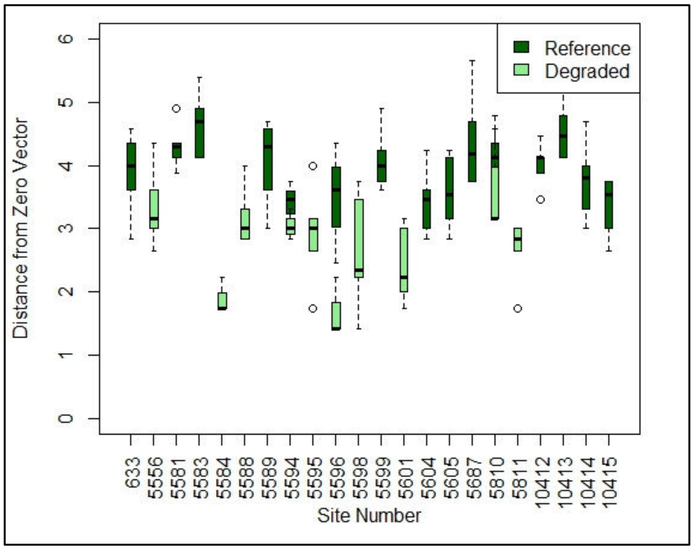

Figure 6: Species richness [mature, sapling, seedling] distance from [0, 0, 0] at each site. Sites shown as both reference and degraded correspond to those on which one side of the creek is considered reference and the opposite side as degraded.

Forest Wetland Affinity

The average wetland status indicator estimate for the reference group was 1.84 and there was no estimated change from the reference group to the degraded group of sites (Table 6). This would indicate that a large number of the species present in each site were upland or FACU species. The average percentage of species that were classified as FAC, FACW, or OBL was approximately 30% for the reference group and 32% for the degraded group. The remainder of the species would be FACU or UPL. Variability was high between sites within each group (Figure 7) and it was necessary to allow for site-specific adjustments within the model. Sites 5581, 5583, 5584, and 5588 had lower estimates of wetland status while sites 5594, 5595, 5596, 5599, 5601, and 5687 had higher estimates of wetland status.

One complicating factor for this metric is the fact that the number of individuals of each species is not represented in the calculation. Thus, a large number of individuals could exist in the relatively few species of FAC, FACW, or OBL classifications but the metric would not show this distinction. Under such a condition, the site might have a much more saturated soil condition than the metric would calculate.

Tree seedling plantings as part of a restoration effort were conducted in degraded sites 5556, 5588, 5594, 5595, and 5598, and in the reference site 5583 prior to collection of data in 2014. This raised the concern that the average wetland status indicator could have been artificially biased by including species that had been planted and which might not be present in the plot otherwise. Both wetland and non-wetland species were included in the planting events and it was difficult to assess the impact of these plantings on the overall proportion of wetland versus non-wetland species present when the data was collected. However, there was not a distinct difference in the average wetland status indicator between mature trees, saplings, and seedlings at the sites where plantings occurred, potentially indicating that planted seedlings that survived at each site were not of a different wetland status from trees that already existed. Thus, it is unlikely that the planting of seedlings at these sites would alter the wetland status indicator for the site as a whole.

Table 6: Model estimates, standard deviation (SD), and credence intervals for the average wetland status indicator. The slope shows no difference between the reference group and the degraded group.

	Estimate	SD	Credence Interval (95%)	
Intercept	1.84	0.05	1.74	1.94
Slope	-0.01	0.04	-0.10	0.07

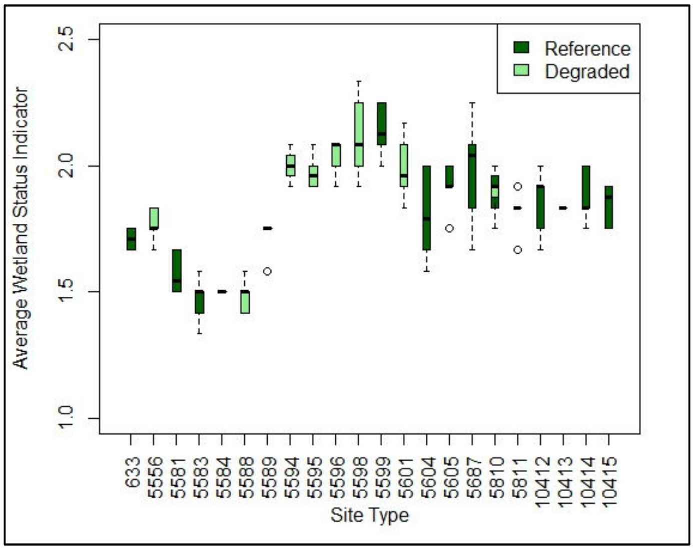

Figure 7: Average wetland status indicator at each site. The wetland status indicator ranged from one for upland species to five for obligate species. Sites shown as both reference and degraded correspond to those on which one side of the creek is considered reference and the opposite side as degraded.

Riparian Zone Width

The riparian zone width and slope estimate was produced for each drainage area to account for variation between drainage areas in both the reference group and the degraded group of sites. There was no pattern to the riparian zone width or slope related to the drainage area size but there was a decrease in slope from the reference group to the degraded group in each drainage area (Table 7). Thus, the riparian zone width was shorter in degraded group regardless of drainage area. Variability was high for reference sites within the 0-64 acre and the 321-640 acre drainage areas as well as degraded sites within the 641-1280 acre drainage area (Figure 8). It was necessary to allow for site specific adjustments within the model especially site 5594 which had small riparian zone widths (Appendix E).

Table 7: Model estimates, standard deviation (SD), and credence intervals for the riparian zone width. Intercepts and slopes are separated by drainage area (1= 0-64, 2=65-350, 3=321-640, 4=641-1280 acres). The slopes show the decrease from the reference group to the degraded group in each drainage area.

	Estimate	SD	Credence Interval (95%)	
Intercept[1]	31.9	4.2	23.8	40.2
Intercept [2]	45.5	4.8	36.0	54.6
Intercept [3]	36.9	3.9	29.1	44.6
Intercept [4]	43.6	3.6	36.5	50.6
Slope[1]	-20.2	5.9	-30.8	-7.9
Slope[2]	-34.9	6.8	-48.0	-22.2
Slope[3]	-27.7	5.9	-39.9	-16.1
Slope[4]	-26.7	4.2	-35.0	-18.5

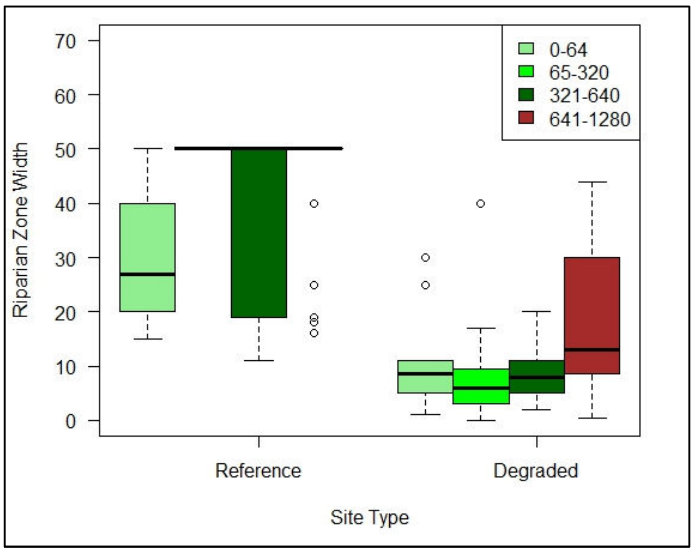

Figure 8: Riparian zone width for drainage area within each site type.

In–stream Cover

The in-stream cover and slope estimate was produced for each drainage area to account for variation between drainage areas in both the reference group and the degraded group of sites. There was no difference in the in-stream cover between the reference group and the degraded group (Table 8) possibly due to the large variation within drainage areas (Figure 9) or within sites (Appendix F). It was necessary to allow for site-specific adjustments within the model, especially site 5584 which had no in-stream cover and site 5687 which had very little in-stream cover. Since in-stream cover does not differentiate cover projected from tall grasses, forbs, or trees, degraded sites with overgrown herbaceous vegetation may have high in-stream cover values. In addition, individual leaves and limbs low in height have a disproportionate effect on the densiometer cover reading when compared to taller trees (Figure 10), potentially inflating the effect of stream cover and the shade it provides in degraded sites from tall herbaceous vegetation, such as giant ragweed (*Ambrosia trifida*), or young trees (Vora 1988).

Table 8: Model estimates, standard deviation (SD), and credence intervals for the in-stream cover. Intercepts and slopes are separated by drainage area (1= 0-64, 2=65-350, 3=321-640, 4=641-1280 acres). The slopes show no difference between the reference group and the degraded group in each drainage area.

	Estimate	SD	Credence Interval (95%)	
Intercept[1]	73.7	12.4	50.5	99.9
Intercept [2]	68.9	11.5	45.5	91.3
Intercept [3]	66.9	11.4	43.1	88.7
Intercept [4]	65.2	11.7	40.1	86.5
Slope[1]	-19.8	16.7	-53.1	14.2
Slope[2]	-15.8	15.5	-45.3	17.2
Slope[3]	-28.8	18.3	-68.5	4.1
Slope[4]	-17.7	15.9	-47.6	16.0

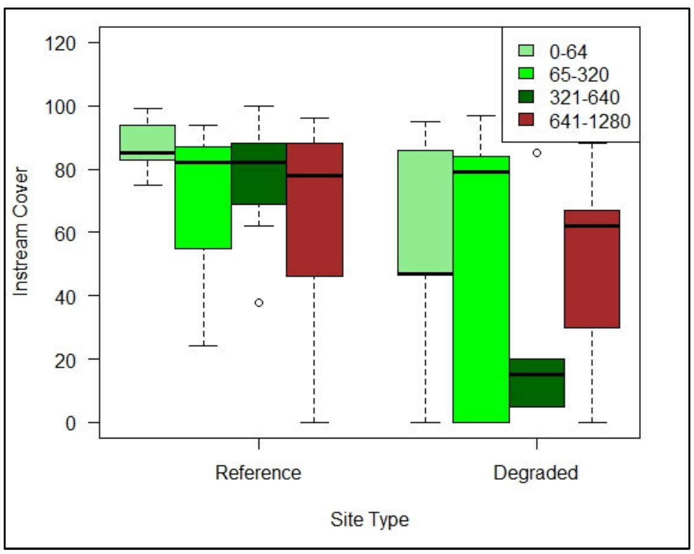

Figure 9: In-stream cover for drainage area within each site type.

Figure 10: Vegetation over the stream (Site 5601, 50 m point along transect). Individual leaves on low branches occupy a larger proportion of the grid than those on higher branches.

Snags

Model estimates for the number of snags at a site ranged from eight to ten within the reference group across drainage areas while the model estimates were significantly less within the degraded group across drainage areas with estimates that ranged from one to two snags in a site (Table 9). Estimates are slightly lower than the means of raw data shown in Figure 11 due to the fact that the higher values of snags will influence the model estimate less than the mean of the raw data. This clearly showed that a higher number of snags were present in the reference group regardless of drainage area. Degraded sites are often in recreational areas with relatively high human recreational use and senescent/dead trees are commonly removed to reduce the risk of injuries caused by fallen limbs or trees. However, management of these areas is shifting to a less drastic approach and focusing on removing only those trees that represent a high or very high likelihood of falling and have a high likelihood of causing a substantial impact on a target which cannot be substantially reduced by other means such as partial tree/limb removal.

Table 9: Model estimates, standard deviation (SD), and credence intervals for the number of snags (log transformed). Intercepts and slopes are separated by drainage area (1= 0-64, 2=65-350, 3=321-640, 4=641-1280 acres). The slopes show the decrease from the reference group to the degraded group in each drainage area.

	Estimate	SD	Credence Interval (95%)	
Intercept[1]	2.30	0.37	1.57	3.12
Intercept [2]	2.08	0.39	1.23	2.82
Intercept [3]	2.32	0.38	1.59	3.15
Intercept [4]	2.27	0.38	1.52	3.05
Slope[1]	-1.69	0.65	-3.12	-0.45
Slope[2]	-1.82	0.66	-3.31	-0.68
Slope[3]	-1.78	0.66	-3.26	-0.57
Slope[4]	-1.69	0.62	-3.06	-0.55

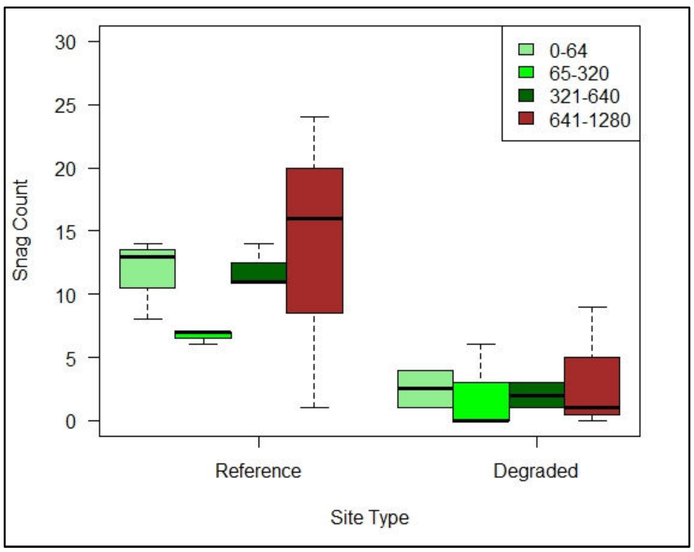

Figure 11: Number of snags for drainage area within each site type.

Large Woody Debris

The model for large woody debris consisted of a general intercept and single slope with intercept adjustments for site and drainage area as opposed to separate intercepts and slopes for each drainage area. Estimates for large woody debris within the reference group ranged from three to fifteen pieces across drainage areas while the counts within the degraded group ranged from one to four across drainage areas (Table 10). The model slope showed a decrease in large woody debris from the reference group to the degraded group regardless of drainage area. Degraded sites, especially those with a chronic history of mowing and few or no trees, have a limited or non-existent source of large woody debris. This is exacerbated by the fact that some degraded sites are managed to maximize stormwater conveyance and reduce flooding risk by removing woody debris pieces that may obstruct culverts. There was no linear pattern for large woody debris across drainage areas (Figure 12, e.g. LWD increasing or decreasing as a function of drainage area. However, the smallest drainage areas had substantially higher counts and the largest drainage areas had substantially lower counts.

Table 10: Model estimates, standard deviation (SD), and credence intervals for the pieces of large woody debris (log transformed). Each lambda is a drainage area specific adjustment to intercept (1= 0-64, 2=65-350, 3=321-640, 4=641-1280 acres). The slope shows the decrease from the reference group to the degraded group.

	Estimate	SD	Credence Interval (95%)	
Intercept	1.84	0.42	1.01	2.66
Slope	-1.46	0.27	-2.02	-0.95
lambda[1]	0.90	0.43	0.06	1.76
lambda[2]	-0.61	0.46	-1.52	0.27
lambda[3]	0.54	0.43	-0.30	1.39
lambda[4]	-0.93	0.47	-1.89	-0.03

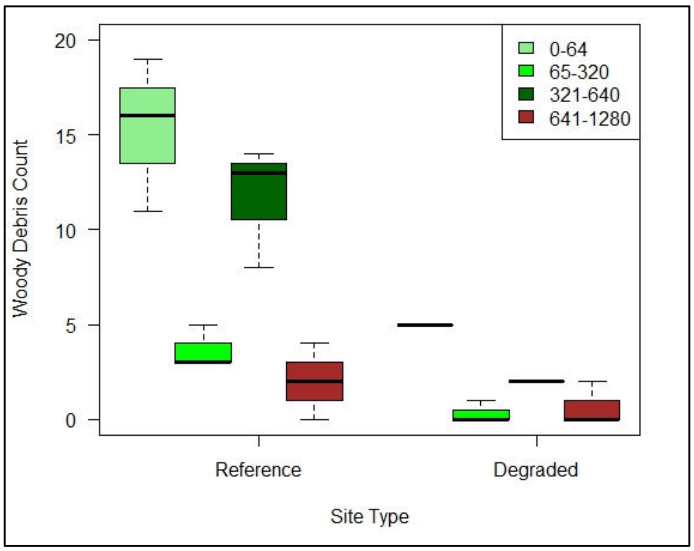

Figure 12: Pieces of large woody debris for drainage area within each site type.

Soil Organic Carbon

There was only one sample of soil organic carbon collected at each site, so no estimate of variation could be done within each site. Models were built to compare the soil organic carbon across sites and drainage areas. Neither site nor drainage area showed a significant impact on the concentration of organic carbon. The model presented strictly shows the comparison of soil organic carbon between the reference group and the degraded group (Table 11). There was no estimated change in soil organic carbon between the two groups with a high variation within the groups (Figure 13).

Table 11: Model estimates, standard deviation (SD), and credence intervals for soil organic carbon. The slope shows no difference between the reference group and the degraded group.

	Estimate	SD	Credence Interval (95%)	
Intercept	4.5	0.6	3.4	5.7
Slope	-0.7	0.9	-2.4	1.0

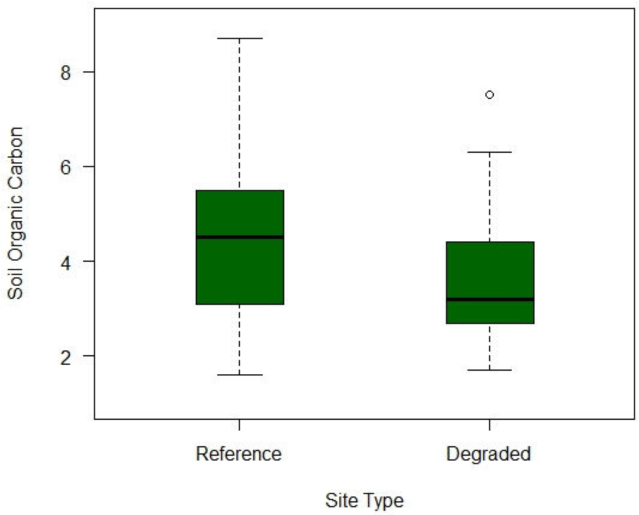

Figure 13: Soil organic carbon for each site type.

Hydrophytic Vegetation

The presence of hydrophytic vegetation differs from many of the other plant metrics because it is measured in sections within the stream channel and not within a plot that begins at bankfull depth. In addition, this metric includes the presence of grasses and forbs in addition to woody species. The probability of occurrence for hydrophytic vegetation was close to one in both the reference group and the degraded group and the difference between the two probabilities included zero (Figure 14). Thus there was no difference in the occurrence of hydrophytic vegetation between the reference group and the degraded group.

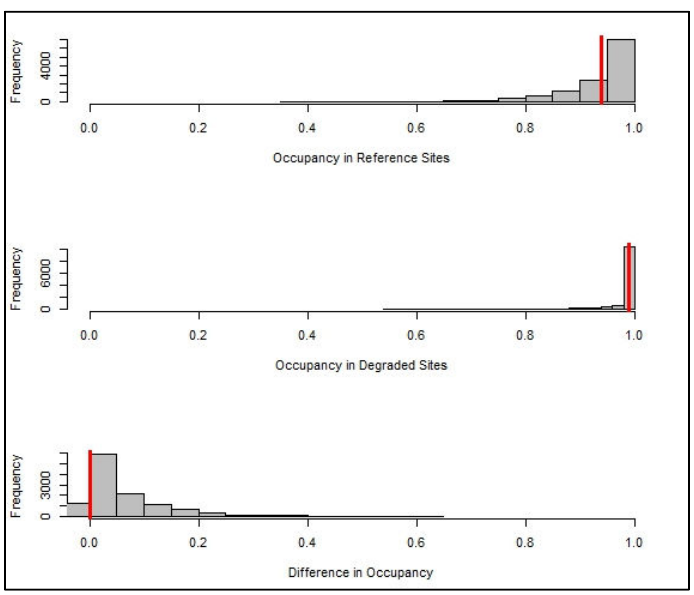

Figure 14: Hydrophytic vegetation occupancy in reference and degraded sites. The difference in occupany is close to zero.

Site Classification and Random Forest Model

The OOB error estimate for the random forest model was 4% which equates to a misclassification of one site from the model. The site that was misclassified was site 5810 which is a site that has been disturbed on only one side of the creek. Observations were pooled across the entire site for the random forest classification model. When the observations were pooled for site 5810, the undisturbed side likely was influential on the site observations input into the random forest model.

The mean decrease in accuracy was high for the species richness distance (siteDistRichness), soil compaction (siteSoilCompaction), riparian zone width (siteRipZoneWidth), plant structural distance (siteDistStructure), total species richness (siteAvgRichness), and the number of snags (siteSnags) (Figure 15). These metrics were deemed important to the classification and were used in further analysis. Other metrics were dropped from further analysis. The metrics determined as important for site classification by the random forest classification are consistent with the results from the Bayesian hierarchical models. A random forest model was performed using only the metrics deemed important to ensure that a loss of information did not occur when the other metrics were dropped. The OOB error estimate for the reduced random forest model was again 4% with the same misclassification of site 5810.

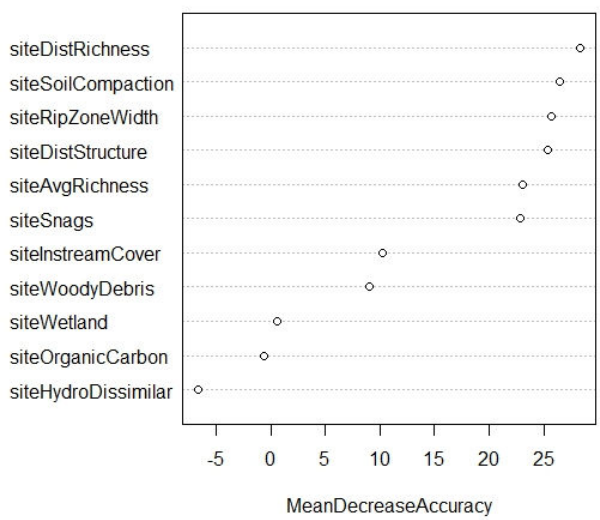

Figure 15: Importance for environmental metrics in the random forest classification model. Large mean decreases in accuracy equate to high importance.

The probability density and mass functions for important metrics to the classification of reference and degraded sites are plotted in Figure 16. With the exception of the number of snags, these functions were based on plot level data within sites. Thus, the sites which contained undisturbed plots on one side of the creek and disturbed plots on the other side of the creek could be separated. This should allow for less error in the classification of a site with the naïve Bayes classifier. Probability functions typically showed that the distribution of values was distributed over lower values in the degraded group with the exception of soil compaction which was distributed over higher values in the degraded group. Probability functions were not disjoint for any environmental metric. Site averages for each metric were computed and inserted as an x-value in the reference functions and the degraded functions.

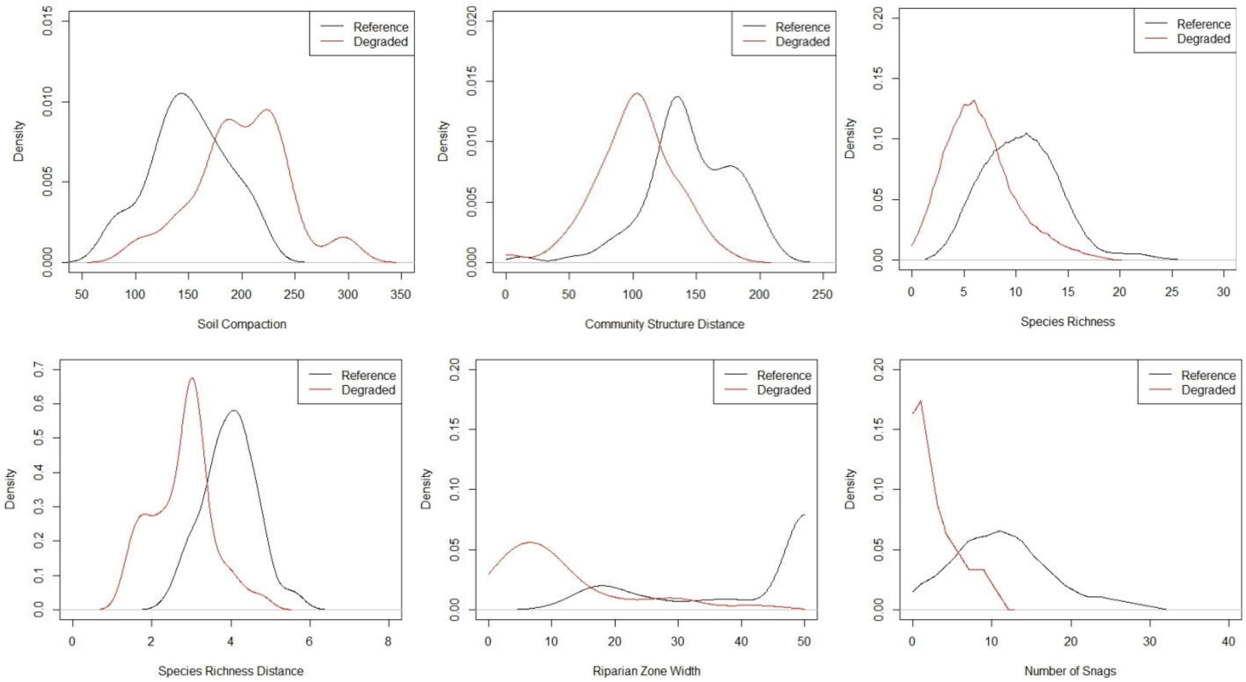

Figure 16: Probability density functions (probability mass functions) for important metrics by degraded or reference site type.

The Hellinger Distance was 0.8165, 0.8851, 0.5289, 0.5438, 0.8907, and 0.8415 for soil compaction, plant structural distance, total species richness, species richness distance, riparian zone width, and number of snags, respectively. For calculation of the RFA index, the riparian zone width was removed due to the management problems associated with this metric. One goal of the RFA is to track the change over time of environmental conditions in degraded sites to see if they become similar to conditions present in the reference sites. In some of the sites, the riparian zone width is currently at a maximum as the site is adjacent to a road or parking lot. As the riparian zone width may not change, it did not make sense to include this metric in the RFA tool. The overall Hellinger Distance was computed as 0.9996 using the following equation:

$$HD = \sqrt{1 - \prod_{i=1}^{5} (1 - Metric\ HD_i^2)}$$

where the *Metric* HD_i was the Hellinger Distance from the five environmental metrics.

The probability that a site belongs to the reference group given site conditions, $P(c_{ref}|\boldsymbol{x})$, the probability that a site to belongs to the degraded group given site conditions, $P(c_{deg}|\boldsymbol{x})$, and the RFA index were then calculated (Table 12, Figure 17). Using the $P(c_{ref}|\boldsymbol{x})$ and $P(c_{deg}|\boldsymbol{x})$, a site would be classified as reference site if $P(c_{ref}|\boldsymbol{x}) > P(c_{deg}|\boldsymbol{x})$. Under these criteria, all of the sites were classified correctly using the naïve Bayes classifier. More importantly, the RFA index is 100 multiplied by the probability of a site to be in the reference group given site conditions with a correction factor to account for distribution shift of the environmental metrics.

Table 12: Calculated $P(c_{ref}|\boldsymbol{x})$, $P(c_{deg}|\boldsymbol{x})$, and RFA index for each site. Shaded cells represent degraded sites.

Site #	Site Name	Watershed	$P(c_{ref}\|\boldsymbol{x})$	$P(c_{deg}\|\boldsymbol{x})$	RFA index
633	Barton Creek Ephemeral 3	Barton	1.00	0.00	100.0
5556	East Bouldin at gabion in Gillis Park	East Bouldin	0.12	0.88	11.9
5581	Bee ds Loop 360	Bee	0.92	0.08	91.7
5583	Blunn us Cow Trough Spring	Blunn	0.91	0.09	91.1
5584	Buttermilk at Buttermilk Park	Buttermilk Branch	0.01	0.99	0.9
5588	Common Ford ds crossing in Common Ford Ranch	Lake Austin	0.13	0.87	12.9
5589	Common Ford us bridge at Common Ford Ranch	Lake Austin	0.84	0.16	83.7
5594	Shoal at Shady Oak Court	Shoal	0.14	0.86	13.9
5595	Tannehill at Bartholomew Park near Berkman Dr.	Tannehill Branch	0.09	0.91	9.5
5596	Tannehill us storm pipe in Givens Park	Tannehill Branch	0.00	1.00	0.4
5598	Taylor Slough South in Reed Park at Footbridge	Taylor Slough South	0.31	0.69	31.1
5599	Little Walnut Trib at Gus Garcia Park	Little Walnut	0.95	0.05	94.5
5601	Walnut Trib at North Star Greenbelt	Walnut	0.04	0.96	3.8
5604	West Bouldin in West Bouldin Greenbelt	West Bouldin	1.00	0.00	100.0
5605	Williamson at Wagon Bend Trl	Williamson	1.00	0.00	100.0
5687	Fort Branch at Tura Ln LISI 1	Fort Branch	1.00	0.00	100.0
5810	Harpers Branch at Parker Park	Harper's Branch	0.39	0.61	38.9
5811	Shoal at Crestmont Park	Shoal	0.18	0.82	17.7
10412	West Bull at Long Canyon and Standing Rock	West Bull	0.91	0.09	91.3
10413	Barton at key west from Cape Coral	Barton	0.91	0.09	90.7
10414	South Boggy at Latteridge Almondsbury	South Boggy	1.00	0.00	100.0
10415	Onion at Nuckols and Thaxton	Onion	0.72	0.28	72.2

Degraded sites and reference sites are clearly segregated based on their RFA index (Figure 17). Over time, as degraded sites ecological function improves, their RFA index is expected to have more similar values to reference sites. As a monitoring tool, the RFA provides a quantitative approach to assess the progress of restoration.

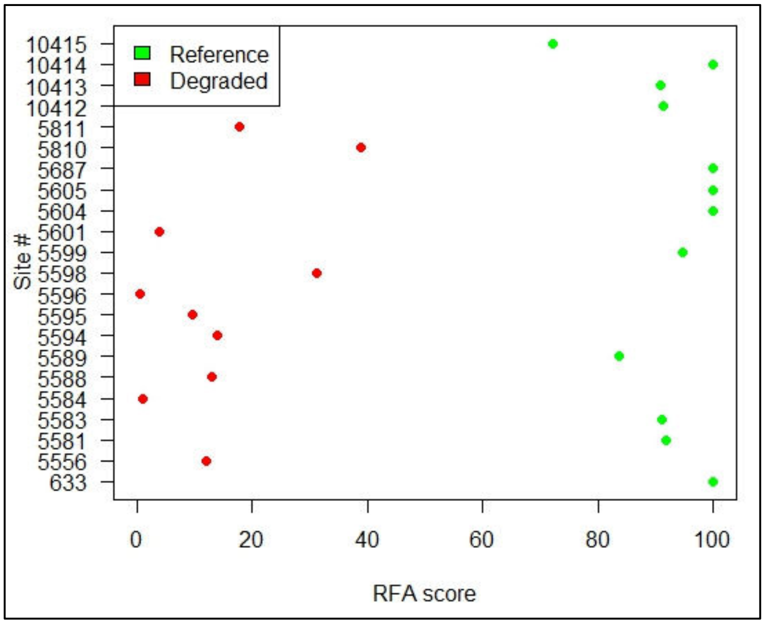

Figure 17: Calculated RFA index for each site. Green points represent reference sites while red points represent degraded sites.

Conclusions and Recommendations

Soil compaction, plant structural distance, total species richness, richness distance, number of snags, large woody debris count, and the riparian zone width were metrics with a quantitative difference between undisturbed urban sites and sites with a history of continuous disturbance but managed for riparian restoration. Five of these metrics were used as inputs to the updated Riparian Functional Assessment index. It is likely that the large woody debris count was too dependent on the drainage area to be considered as an important variable in the random forest classification model. In addition, the riparian zone width was deliberately excluded from the RFA index because it is unlikely that this metric will change at many restoration sites given management constrains in parks and proximity to infrastructure such as parking lots present in some sites. As degraded riparian zones are left undisturbed, the soil should become less compacted and the plant communities should undergo a successional process. Thus, each of the other metrics are expected to shift over time. It is recommended that these seven metrics continue to be collected at each site even though they are not all included in the RFA index because they help monitor change about specific functional elements of riparian areas and guide whether or not management actions aimed at restoring them are needed. For example, large woody debris provides habitat for aquatic organisms and restoration practices can consider mechanisms to increase its presence whenever it does not conflict with flooding prevention.

The average wetland status indicator, in-stream cover, soil organic carbon, and presence of hydrophytic vegetation were not shown to be different between undisturbed urban sites and disturbed urban sites under riparian restoration management. Average wetland status could be improved if the calculation was a weighted average instead of a normal average where the weights were determined by the number of individuals of each species. Under current sampling, it is impossible to distinguish the dominance of a species of tree and thus each species receives an equal weight. However, given the logistic constraints of counting individuals for each species, we do not recommend a methodological change for vegetation sampling. Although additional bias could be introduced to the average wetland status by plantings that occur in the degraded sites as part of the riparian restoration efforts, the seedling plantings that have taken place in some of the degraded sites included a mix of plants along a moisture gradient requirement. Current analysis indicated that planted seedlings that survived at each site were not of a different wetland status from trees that already existed and it is unlikely that the planting of seedlings at these sites would alter the average wetland status.

It is also possible, but unlikely, that the seedling plantings could have impacted the species richness at a degraded site. Seedling survival is very low and has been shown not to significantly increase species diversity after a year of planting (González and Richter 2014). In addition, species richness was not higher in the seedling layer when compared to the sapling layer for sites where planting events had been conducted. Thus, similar or higher levels of diversity were already present at these sites in age classes present prior to the planting events.

In-stream cover was highly variable within a site for both reference and degraded sites. A different methodology may more accurately represent the in-stream cover at a site. Aerial photography could be used to determine a percentage of the stream that is covered by vegetation on a given year. However, the frequency of aerial photography generation may not allow for updates on each of the riparian zones every other year. In addition, aerial photography is often taken during the leaf-off months and image classification methods may produce canopy estimates that miss deciduous trees often associated with riparian forests. A more feasible methodology change would be to visually estimate the in-stream canopy cover and/or directly measure the light intensity at the same location where the densiometer reading of canopy cover is collected. It is recommended to alter the RFA sampling collection to include a visual estimate of in-stream canopy cover, densitometer reading, and light intensity measurement at each sample location to determine if there is a difference in the data collected from each of these methods.

The RFA index has minimal direct inputs for soil characteristics. Compaction is the only soil characteristic included directly into the tool. However, soil characteristics should change as plant communities establish within degraded sites. Furthermore, riparian soils provide critical water quality functions and it is important to monitor improvements in its function in degraded sites. Further study is needed to fully understand exactly how the soil characteristics will change in urban riparian zones. It is recommended that further data collection of soil characteristics within urban reference sites and urban degraded sites be included in the RFA project but the data will not be included into the index. More specifically, the soil texture, soil structure, soil nutrients, active bacteria, and active fungi should be analyzed in future data collection. A study over a short period of time to examine the spatial variation in soil parameters within a site is recommended to inform if aggregating soil samples into a composite is a good practice or if it should be modified.

The infiltration rate of riparian soils is influenced by both soil texture as well as soil structure. Soil texture is a physical characteristic inherent to the soil and determined by distribution of particle sizes. Soil texture is often considered a basic property of the soil because changes in soil texture occur over large temporal scales (Pierzynski et al. 1994). Soil structure, developed by the formation of aggregates in the soil, is dynamic and altered by physical processes such as wetting/drying and freezing/thawing cycles, root extension, the action of soil organisms, tillage and other disturbances, and organic matter accumulation (Brady 1990). Soil disturbance such as tillage has long-term deleterious effects on soil structural features that control water infiltration and subsequent water transmission and storage in soil (Franzluebbers 2002). Over time, as vegetation develops in degraded sites where mowing and disturbance has been mostly removed, soil structure is likely to improve with an associated enhancement of water infiltration. Water infiltration is a critical function of riparian areas for improved water quality, thus, we recommend measuring the infiltration rate of these sites relative to reference sites utilizing a double-ring infiltrometer or mini-disk infiltrometer to minimize disturbing the soil structure.

The presence of hydrophytic vegetation below bankfull depth was not different between reference and degraded sites because the vegetation was present in all sites. It is recommended to count the number of species as well as measure the percent cover of hydrophytic vegetation within a 0.5 m of bankfull depth into the riparian area. This may help us determine a difference in the hydrophytic plant community present at degraded and reference sites since this work has shown that a community exists within each site type.

References

Abbe, TB and DR Montgomery. 1996. Large woody debris jams, channel hydraulics and habitat formation in large rivers. Regulated Rivers Research & Management 12: 201-221.

Alpaydin, E. 2010. Introduction to Machine Learning, Second Edition. The MIT Press, Cambridge, Massachusetts. 47 – 60 pp.

Anderson, NH and JR Sedell. 1979. Detritus processing by macroinvertebrates in stream ecosystems. Annual Review of Entomology 24: 351-377.

Barbour, MT, J Gerritsen, BD Snyder, and JB Stribling. 1999. Rapid Bioassessment Protocols for use in streams and wadeable rivers, Second Edition. EPA 841-B-99-002. U.S. Environmental Protection Agency; Office of Water; Washington, D.C.

Basu, A, IR Harris, and S Basu. 1997. Minimum distance estimation: The approach using density-based distances. In *Handbook of Statistics*, 15: 21–48.

Brady, NC. 1990. The Nature and Properties of Soils, 10th ed. Prentice Hall, New Jersey, USA.

Breiman, L. 2001. Random forests. Machine Learning 45(1): 5-32.

Breiman, L. 2002. Manual on setting up, using, and understanding random forests v3.1, http://oz.berkeley.edu/users/breiman/Using_random_forests_V3.1.pdf.

Correll, D. 1999. Vegetated Stream Riparian Zones Their Effects on Stream Nutrients, Sediments, and Toxic Substances: An Annotated and Indexed Bibliography of the world literature including buffer strips, and interactions with hyporheic zones and floodplains, 8th Ed. Smithsonian Environmental Research Center Edgewater, Maryland, USA.

Duncan, A, S Wagner, M Scoggins, A Richter. 2011. Riparian reference condition: Using regional plant composition to guide functional improvements in the City of Austin. City of Austin, Watershed Protection Department, Environmental Resource Management, SR-11-13.

Duncan, A. 2012. A functional approach to riparian restoration in Austin, Texas. City of Austin, Watershed Protection Department, Environmental Resource Management, SR-12-05.

Fischer, RA and JC Fischenich. 2000. Design recommendations for riparian corridors and vegetated buffer strips. EMRRP Technical Notes Collection (ERDC TN-EMRRP-SR-24), U.S. Army Engineer Research and Development Center, Vicksburg, MS.

Fowler, NL and MT Simmons. 2008. Savanna dynamics in central Texas: just succession? Applied Vegetation Science 12: 23-31.

Franzluebbers, AJ. 2002. Water infiltration and soil structure related to organic matter and its stratification with depth. Soil and Tillage Research 66(2): 197-205.

Gift, DM, PM Groffman, SS Kaushal, and PM Mayer. 2010. Denitrification Potential, Root Biomass, and Organic Matter in Degraded and Restored Urban Riparian Zones. Restoration Ecology 18(1): 113–120.

González, A and FA Richter. 2014. Sapling Survival Assessment: Examining Success of One Urban Riparian Restoration Tool in Austin, Texas. City of Austin, Watershed Protection Department, Environmental Resource Management, SR-14-13.

González, A and FA Richter. 2014. Riparian Functional Assessment Phase Two Summary. City of Austin, Watershed Protection Department, Environmental Resource Management, SR-14-18.

Groffman, PM, DJ Bain, LE Band, KT Belt, GS Brush, JM Grove, RV Pouyat, IC Yesilonis, and WC Zipperer. 2003. Down by the riverside: urban riparian ecology. Frontiers in Ecology and Environment 1(6): 315–321.

Gurnell, A, K Tockner, P Edwards, and G Petts. 2005. Effects of deposited wood on biocomplexity of river corridors. Frontiers in Ecology and the Environment 3(7): 377-382.

Harmon, ME, JF Franklin, FJ Swanson, P Sollins, and SV Gregory. 1986. Ecology of coarse woody debris in temperate ecosystems. Advances in Ecological Research 15: 133-302.

Hobbs, R J and VA Cramer. 2008. Restoration Ecology: Interventionist Approaches for Restoring and Maintaining Ecosystem Function in the Face of Rapid Environmental Change. Annual Review of Environment And Resources 33: 39-61.

Hyatt, TL and RJ Naiman. 2001. The Residence Time of Large Woody Debris in the Queets River, Washington, USA. Ecological Applications 11: 191-202.

Jacobs, TC and JW Gilliam. 1985. Riparian losses of nitrate from agricultural drainage waters. Journal of Environmental Quality 14: 472-478.

Kailath, T. 1967. The Divergence and Bhattacharyya Distance Measures in Signal Selection. IEEE Transactions on Communications, 15(1): 52–60.

Karr, JR and EW Chu. 1997. Biological Monitoring and Assessment: Using Multimetric Indexes Effectively. EPA 235-R97-001. University of Washington, Seattle.

Larson, MG, DB Booth, and SA Morley. 2001. Effectiveness of large woody debris in stream rehabilitation projects in urban basins. Ecological Engineering 18: 211-226.

Lichvar, RW and JT Kartesz. 2009. North American Digital Flora: National Wetland Plant List, version 2.4.0 (https://wetland_plants.usace.army.mil). US Army Corps of Engineers, Engineer Research and Development Center, Cold Regions Research and Engineering Laboratory, Hanover, NH, and BONAP, Chapel Hill, NC.

Lowrance, R, RL Todd, and LE Asmussen. 1983. Waterborne nutrient budgets for the riparian zone of an agricultural watershed. Agriculture, Ecosystems & Environment 10: 371-384.

Lowrance, R, R Leonard, and J Sheridan. 1985. Managing riparian ecosystems to control nonpoint pollution. Journal of Soil & Water Conservation 40: 87-97.

Magurran, AE. 1988. Ecological Diversity and its Measurement. Princeton University Press, Princeton, U.S.A.

Moffatt, SF, SM McLachlan, and NC Kenkel. 2004. Impacts of Land Use on Riparian Forest along an Urban-Rural Gradient in Southern Manitoba. Plant Ecology 174(1): 119-135.

Naiman, RJ and H Decamps. 1997. The Ecology of Interfaces: Riparian Zones. Annual Review of Ecology and Systematics 28: 621-658.

von der Ohe, PC, A Prub, RB Schafer, M Liess, E Deckere De, and W Brack. 2007. Water quality indices across Europe – a comparison of the good ecological status of the five river basins. Journal of Environmental Monitoring 9: 970–978.

Peterjohn, WT and DL Correll. 1984. Nutrient dynamics in an agricultural watershed: Observations on the role of a riparian forest. Ecology 65: 1466-1475.

Pierzynski, GM, JT Sims, and GF Vance. 1994. Our Environment: Atmosphere, Hydrosphere, Soils. In: Soils and Environmental Quality. Lewis Publishers, Boca Raton, FL. 7-39 pp.

Richardson, MD, PM Holmes, KJ Esler, SM Galatowitsch, JC Stromberg, SP Kirkman, P Pysek, and RJ Hobbs. 2007. Riparian vegetation: degradation, alien plant invasion, and restoration prospects. Diversity and Distributions 13: 126-139.

Richter, FA and A Duncan. 2012. Riparian Functional Assessment: Choosing Metrics that Quantify Restoration Success in Austin, TX. City of Austin, Watershed Protection Departement, Environmental Resource Management, SR-12-12.

Riseng, CM, MJ Wiley, RJ Stevenson, TG Zorn, and PW Seelbach. 2006. Comparison of Course versus fine scale sampling on statistical modeling of landscape effects and assessment of fish assemblages of the Muskegon River, Michigan. American Fisheries Society Symposium 48: 555–575.

Stacey, PB, A Jones, JC Catlin, DA Duff, LE Stevens, and C Gourley. 2006. User's Guide for the Rapid Assessment of the Functional Condition of Stream Riparian Ecosystems in the American Southwest. Wild Utah Project.

Stromberg, JC, R Tiller, and B Richter. 1996. Effects of Groundwater Decline on Riparian Vegetation of Semiarid Regions: The San Pedro, Arizona. Ecological Applications 6: 113-131.

Sung, CY, L Ming-Han, GO Rogers, A Volder, and Z Wang. 2011. Investigating alien plant invasion in urban riparian forests in a hot and semi-arid region. Landscape and Urban Planning 100: 278–286.

Suter, GW. 1993. A critique of ecosystem health concepts and indexes. Environmental Toxicology and Chemistry 12: 1533–1539.

USDA National Resources Conservation Service. 2008. Soil Quality indicators: Bulk Density. http://soils.usda.gov/sqi/assessment/files/bulk density sq physical indicator sheet.pdf

Vora, RS. 1988. A comparison of the spherical densiometer and ocular methods of estimating canopy cover. Western North American Naturalist 48: 224-227.

Wissmar, RC and RL Beschta. 1998. Restoration and management of riparian ecosystems: a catchment perspective. Freshwater Biology 40: 571-585.

Woolsey, S, F Capelli, T Gonser, E Hoehn, M Hostmann, B Junker, A Paetzold, C Roulier, S Schweizer, SD Tiegs, K Tockner, C Webber, and A Peter. 2007. A strategy to assess river restoration success. Freshwater Biology 52(4): 752-769.

Woolsey, S, C Weber, T Gonser, E Hoehn, M Hostmann, B Junker, C Roulier, S Schweizer, S Tiegs, K Tockner, and A Peter. 2005. Handbook for evaluating rehabilitation projects in rivers and streams. Publication by the Rhone-Thur project. Eawag, WSL, LCH-EPFL, VAW ETHZ. 108 pp.

Appendix A: Results for the t-test for soil compaction. Alpha is a general intercept for the soil compaction, beta is a general difference in soil compaction between reference and degraded sites, while delta is a site specific intercept adjustment.

	Estimate	SD	Credence Interval (95%)	
alpha	146.8816	5.121176	136.1	156.4525
beta	51.48061	7.727817	36.62425	66.75525
delta[633]	4.636716	8.971217	-10.4605	26.7505
delta[5556]	-2.98471	8.527641	-22.9615	12.6415
delta[5581]	1.517914	8.251564	-14.64	20.33575
delta[5583]	1.538433	8.097851	-14.6058	20.0415
delta[5584]	0.991249	8.165346	-15.79	19.12725
delta[5588]	3.844966	8.622292	-11.5763	24.47625
delta[5589]	-7.33783	9.591292	-29.5105	6.45615
delta[5594]	-0.83497	8.119855	-18.24	16.10525
delta[5595]	-1.82458	8.276169	-20.4953	14.50525
delta[5596]	2.633928	8.186797	-13.0205	21.55575
delta[5598]	-1.83529	8.178142	-20.5953	14.69575
delta[5599]	8.018438	10.20793	-5.91655	33.0625
delta[5601]	5.124181	9.03616	-9.5047	26.78675
delta[5604]	1.508379	8.366417	-15.0183	20.881
delta[5605]	-6.83415	9.413673	-29.3553	7.23275
delta[5687]	7.236006	9.959639	-7.37743	32.22525
delta[5810]	1.813314	8.097043	-14.2553	20.26525
delta[5811]	-6.38877	10.06089	-31.4568	8.88685
delta[10412]	2.92625	8.671112	-13.3968	23.142
delta[10413]	-1.22216	7.905824	-18.921	15.34625
delta[10414]	6.473043	9.74022	-8.44928	29.85525
delta[10415]	-2.42354	8.310287	-21.5408	13.38

Appendix B: Results for the t-test for plant structural distance from [0,0,0] where the vector consists of percent cover in [canopy, understory, groundcover]. Alpha is a general intercept for the plant structural distance, beta is a general difference in plant structural distance between reference and degraded sites, while delta is a site specific intercept adjustment.

	Estimate	SD	Credence Interval (95%)	
alpha	137.7445	5.880181	125.2	148.5
beta	-32.8517	8.36036	-48.4	-15.2648
delta[633]	13.79931	11.58434	-6.5765	38.092
delta[5556]	-7.29797	11.19632	-29.93	14.20125
delta[5581]	2.162117	10.69321	-18.8378	23.83
delta[5583]	10.49767	11.28612	-10.1825	33.38575
delta[5584]	-24.8899	13.1005	-52.1715	-0.43595
delta[5588]	7.666614	11.48245	-13.89	30.8805
delta[5589]	21.05804	12.36399	-1.45558	46.40525
delta[5594]	-5.19474	10.56314	-26.43	15.662
delta[5595]	-5.55136	11.29051	-28.871	16.43625
delta[5596]	-21.8484	11.86698	-45.455	0.221017
delta[5598]	14.98536	11.62496	-6.34165	38.571
delta[5599]	24.05635	13.0802	0.409345	51.71525
delta[5601]	-1.98568	10.99532	-24.351	19.65
delta[5604]	11.17767	11.36477	-9.5202	35.922
delta[5605]	6.918682	11.03529	-13.8958	29.9515
delta[5687]	-12.903	10.99615	-35.0458	7.913725
delta[5810]	4.622495	10.58429	-15.7173	26.38725
delta[5811]	7.226054	11.34717	-14.8858	29.773
delta[10412]	-1.70084	11.13787	-23.2563	21.41825
delta[10413]	-11.2864	10.81987	-32.8283	9.229575
delta[10414]	1.450752	10.76829	-19.4825	23.01525
delta[10415]	11.22569	11.3541	-9.8053	34.9505

Appendix C: Results for the t-test for woody species richness. Alpha is a general intercept for the species richness, beta is a general difference in species richness between reference and degraded sites, while delta is a site specific intercept adjustment.

	Estimate	SD	Credence Interval (95%)	
alpha	10.40273	0.588892	9.265	11.62
beta	-3.75563	0.838933	-5.35313	-2.0919
delta[633]	0.148403	1.098982	-2.028	2.315
delta[5556]	1.205572	1.035446	-0.78821	3.279
delta[5581]	1.035265	1.153269	-1.16208	3.46405
delta[5583]	2.956215	1.274261	0.58605	5.588
delta[5584]	-1.76451	1.21234	-4.15405	0.572815
delta[5588]	0.668152	1.011258	-1.291	2.687
delta[5589]	0.113522	1.11835	-2.06305	2.3611
delta[5594]	-0.82302	0.998677	-2.76003	1.126
delta[5595]	1.00608	1.043891	-0.9789	3.111025
delta[5596]	-1.61957	1.015192	-3.61403	0.386418
delta[5598]	-0.49144	1.003386	-2.506	1.46205
delta[5599]	0.232103	1.10864	-1.86013	2.476
delta[5601]	-1.03292	0.998587	-3.01508	0.9252
delta[5604]	-1.06391	1.104958	-3.26003	1.11905
delta[5605]	-1.047	1.116129	-3.243	1.135025
delta[5687]	1.255649	1.16923	-0.89519	3.7071
delta[5810]	1.224276	1.071769	-0.79882	3.452225
delta[5811]	-0.91259	1.045974	-2.97103	1.165025
delta[10412]	0.094587	1.147386	-2.0601	2.398225
delta[10413]	1.872639	1.208557	-0.33542	4.396025
delta[10414]	-0.86047	1.102688	-3.054	1.34
delta[10415]	-2.47788	1.146009	-4.72115	-0.22977

Appendix D: Results for the t-test for woody species richness distance from [0,0,0] where the vector consists of species richness in [canopy, understory, groundcover]. Alpha is a general intercept for the species richness distance, beta is a general difference in species richness distance between reference and degraded sites, while delta is a site specific intercept adjustment.

	Estimate	SD	Credence Interval (95%)	
alpha	3.920646	0.127003	3.66	4.166
beta	-1.10269	0.184159	-1.456	-0.7267
delta[633]	-0.01703	0.228867	-0.46431	0.435305
delta[5556]	0.348327	0.240958	-0.1167	0.835905
delta[5581]	0.275043	0.236893	-0.18294	0.756003
delta[5583]	0.513629	0.247224	0.052258	1.018
delta[5584]	-0.50368	0.305254	-1.13703	0.054817
delta[5588]	0.23525	0.239004	-0.2278	0.716305
delta[5589]	0.111427	0.233668	-0.33632	0.574338
delta[5594]	-0.10103	0.223936	-0.54522	0.3339
delta[5595]	0.073784	0.238256	-0.39801	0.543005
delta[5596]	-0.54977	0.239369	-1.032	-0.08797
delta[5598]	-0.15868	0.238707	-0.63365	0.295508
delta[5599]	0.112424	0.229465	-0.33281	0.5743
delta[5601]	-0.29978	0.24448	-0.7968	0.164005
delta[5604]	-0.33439	0.229806	-0.80191	0.106908
delta[5605]	-0.24189	0.22934	-0.702	0.206508
delta[5687]	0.315703	0.237469	-0.13811	0.8011
delta[5810]	0.317344	0.230043	-0.12131	0.77311
delta[5811]	-0.11843	0.247124	-0.60931	0.351808
delta[10412]	0.064483	0.242124	-0.40733	0.542603
delta[10413]	0.471251	0.245405	0.005841	0.961425
delta[10414]	-0.10125	0.228083	-0.55372	0.351203
delta[10415]	-0.38181	0.23407	-0.84952	0.065868

Appendix E: Model results for riparian zone width. Alphas represent drainage area specific intercepts (1= 0-64, 2=65-350, 3=321-640, 4=641-1280 acres), betas are general differences in riparian zone width for each drainage area, while delta is a site specific intercept adjustment.

	Estimate	SD	Credence Interval (95%)	
alpha[1]	31.93072	4.194809	23.78	40.20025
alpha[2]	45.4631	4.804169	36.04	54.63
alpha[3]	36.89501	3.948187	29.06	44.61
alpha[4]	43.64331	3.575213	36.52	50.55
beta[1]	-20.2008	5.894937	-30.76	-7.94195
beta[2]	-34.8764	6.786186	-48.0103	-22.22
beta[3]	-27.7195	5.926633	-39.8608	-16.09
beta[4]	-26.7494	4.221201	-35.0203	-18.51
delta[633]	-2.29641	4.655932	-11.5503	6.652075
delta[5556]	-3.74954	5.007959	-14.19	5.726
delta[5581]	3.433651	5.067179	-6.02015	13.92025
delta[5583]	3.412682	5.080634	-6.03213	13.94
delta[5584]	0.521418	4.902408	-9.23003	10.1605
delta[5588]	8.576024	4.929034	-0.60361	18.85025
delta[5589]	-0.26118	4.372249	-8.87723	8.457025
delta[5594]	-12.176	4.634698	-21.5203	-3.284
delta[5595]	-0.51585	5.176602	-10.8503	9.7271
delta[5596]	-1.23083	4.282769	-9.69003	7.20105
delta[5598]	-1.66831	4.890733	-11.7703	7.785275
delta[5599]	6.530102	4.666154	-2.26828	16.00025
delta[5601]	-4.39644	5.421977	-15.6005	5.763225
delta[5604]	4.628515	4.489624	-3.62025	14.05025
delta[5605]	1.186442	4.651147	-8.07033	10.41
delta[5687]	4.644872	4.431625	-3.62608	13.72
delta[5810]	4.152963	4.526161	-4.69805	13.19
delta[5811]	-0.44023	5.128652	-10.6903	9.712675
delta[10412]	-0.03786	4.61037	-8.948	9.086075
delta[10413]	-3.98732	4.705649	-13.7703	4.840125
delta[10414]	-6.96212	4.629688	-16.37	1.818275
delta[10415]	3.410915	5.087894	-6.0253	13.91025

Appendix F: Model results for instream canopy cover. Alphas represent drainage area specific intercepts (1= 0-64, 2=65-350, 3=321-640, 4=641-1280 acres), betas are general differences in instream canopy cover for each drainage area, while delta is a site specific intercept adjustment.

	Estimate	SD	Credence Interval (95%)	
alpha[1]	73.66899	12.42887	50.54975	99.8605
alpha[2]	68.89451	11.47976	45.46975	91.33025
alpha[3]	66.94368	11.41885	43.06875	88.65125
alpha[4]	65.23379	11.67022	40.1095	86.5305
beta[1]	-19.7548	16.7255	-53.0503	14.2205
beta[2]	-15.7703	15.50842	-45.25	17.2305
beta[3]	-28.7605	18.33469	-68.4805	4.131075
beta[4]	-17.6873	15.92773	-47.63	16.04025
delta[633]	5.039912	15.0191	-24.49	34.20075
delta[5556]	23.88212	15.99224	-6.7471	55.88125
delta[5581]	3.64149	14.48676	-24.3813	32.77025
delta[5583]	15.39133	14.74889	-13.0005	45.18
delta[5584]	-42.3045	16.34297	-75.7908	-11.42
delta[5588]	24.47719	15.71048	-5.49128	56.28025
delta[5589]	17.21186	14.67098	-10.57	47.29025
delta[5594]	-0.76758	15.63732	-31.74	30.5005
delta[5595]	-23.8004	17.68449	-57.9708	11.31025
delta[5596]	-20.0094	15.88079	-51.3308	11.09025
delta[5598]	25.69043	15.89254	-4.93138	56.821
delta[5599]	22.33934	14.54238	-5.79308	51.78175
delta[5601]	5.106391	17.2464	-28.9903	39.34
delta[5604]	12.51812	14.71856	-15.8503	42.62
delta[5605]	8.948513	15.08496	-20.3405	39.21025
delta[5687]	-36.7734	14.87954	-65.7803	-7.5438
delta[5810]	-5.5049	17.30476	-39.5705	28.95075
delta[5811]	1.411168	17.62754	-32.2223	38.38025
delta[10412]	7.2835	14.54452	-20.6308	36.7705
delta[10413]	18.05821	15.16264	-11.7703	48.29
delta[10414]	-5.64273	14.32834	-34.03	22.07
delta[10415]	-14.8457	14.58391	-43.68	13.6705