

Developing A Stream Stability Index

SR-15-09, July 2015

Stephen Davis
Aaron Richter, EIT
Mateo Scoggins

City of Austin
Watershed Protection Department
Environmental Resources Management Division

Abstract

Stream channel instability occurs when altered hydrology affects rates of sediment deposition and exacerbates erosion of stream banks. A Stream Stability Index based on effective indicators was desired to identify and prioritize watershed-scale restoration needs as well as help contextualize the results of the City of Austin Watershed Protection Department (WPD) biological monitoring program. The WPD assessed geomorphic, hydrologic, and riparian characteristics in Austin watersheds to determine effective indicators of stream channel instability. A total of 15 parameters were evaluated at 21 stream reaches. Each reach was classified as stable or unstable based on historical data and professional judgement by City scientists. The 15 parameters were chosen because they have been used in long-term routine monitoring physical integrity assessments by WPD. Random forest analysis was used to determine the parameters that most accurately classified both stable and unstable reaches. Identified parameters were then compiled and considered for use in the Stream Stability Index. Results of the model analysis indicated that Scouring and Deposition was the best performing parameter, consistently classifying all site locations accurately. Other parameters that performed well with minimal rates of error were particle size (D_{50}), Consolidation of Particles, and Mass Wasting. In the Edwards Plateau, Obstructions and Deflectors and Bank Protection from Vegetation were important parameters, while Bank Material was an important parameter in the Blackland Prairie. These parameters are recommended for the development of a Stream Stability Index specific to each ecoregion. Models generated in this study utilizing these specific sets of parameters were able to correctly classify all stable and unstable sites in their respective ecoregions.

Introduction

Stream morphology responds to watershed stressors of varying types, duration, and magnitude, which may result in degraded surface water quality, damaged structures, loss of amenity values, and diminished habitat (Hooke 1986, Knox 2006, Rosgen 1996). Morphologic responses to stressors are long term, generally with a marked initial response, and a long-term adjustment period (Knighton 1998). While short-term responses are often easy to monitor visually, long-term responses are not as evident and require detailed data analysis.

Both short term and long term monitoring of stream channel characteristics are necessary for proper watershed management (Chin 2006), and can be achieved through the development and implementation of a Stream Stability Index. A stability index would provide Austin's watershed managers with the ability to assess a streams' current geomorphic condition and compare changes in condition over time. Consistent evaluation of hydrologic, geomorphic, and riparian parameters of stream channels allows for the identification of the areas with major physical degradation and high rates of morphologic change, and for efficient resource allocation to those areas.

Stream stability is a relative term among fluvial geomorphologists. Actual persistent stability never exists in natural streams and rivers. Fluvial systems often change their position and generally fluctuate around a range of sediment loads and discharges. Fluvial systems can become reasonably stable in the sense that they are resilient and resistant to disturbance, and will return to their approximate previous state after disturbance (i.e. stable equilibria, Knighton 1998). The determination of the stability condition of a channel reach along a gradient of relative equilibria is the goal of this stream stability index.

Measurement of the adjustments to channel geometry of fluvial systems involve numerous variables, and the interdependence of these variables is not always clear because the role of a single variable cannot easily be isolated (Knighton 1998). Therefore, this project assessed a range of parameters associated with fluvial systems to determine which ones are predictable indicators of stream stability for Austin, TX.

Methods

Site selection

Twenty one total reaches (selected from City of Austin monitoring locations) were assessed in the Edwards Plateau and Blackland Prairie ecoregions of Austin, TX (Figure 1, Omernik 1987). Ten of the 21 reaches were located in the Edwards Plateau (5 stable and 5 unstable sites, Table 1). The other 11 reaches were in the Blackland Prairie (5 stable and 6 unstable sites, Table 2). The City of Austin spans both ecoregions which differ geologically and topographically. Those differences can lead to varying channel adjustment response times and therefore study reaches in both ecoregions were evaluated. Selection of stable and unstable reach locations was based on development level, channel and bank material, historical monitoring data from Environmental Integrity Index scores (EII) scores, drainage area, and best professional judgement by WPD staff scientists and engineers (COA 2002). In general, stable reaches were selected that had catchment impervious cover levels less than 15%, exhibited consistently higher EII scores (>75 percentile), and had bank and substrate characteristics considered stable by WPD staff. Unstable reaches had higher levels of development and/or had a long history of agricultural degradation, and scored lower in the EII (<50). Among all study reaches, a range of drainage areas was selected that was representative of the variation in monitoring location catchment areas used by WPD long-term monitoring programs throughout the Austin area.

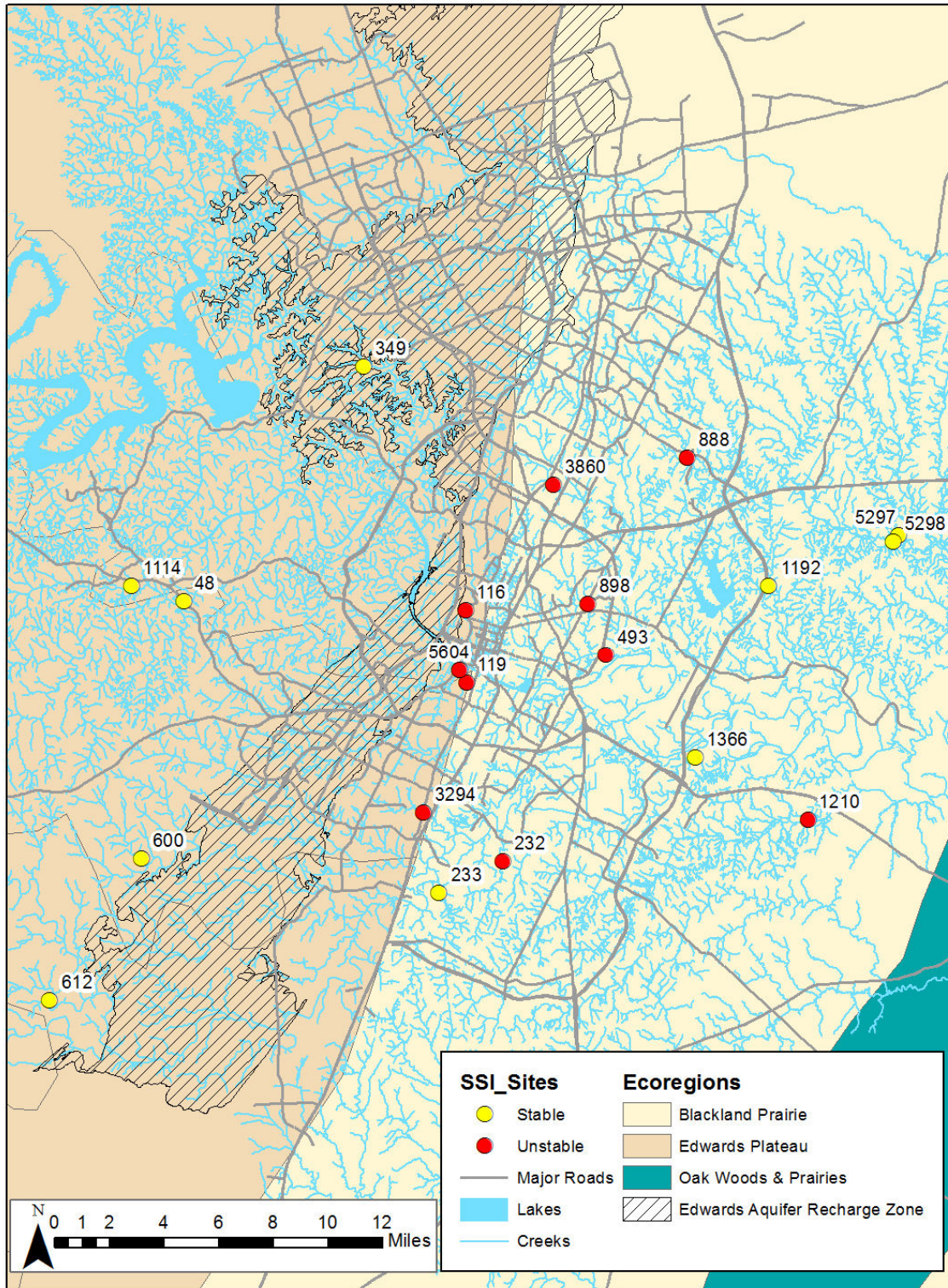


Figure 1. Map of all stable and unstable reference sites located in the Edwards Plateau and Blackland Prairie ecoregions of Austin, TX that were assessed over the duration of this study. Listings of stable and unstable sites and their representative geographic regions can be found in Table 1 and 2.

Table 1. Stable and unstable site locations assessed in the Edwards Plateau ecoregion of Austin, TX.

Edwards Plateau Stable			Edwards Plateau Unstable		
Site No.	Site Name	Watershed	Site No.	Site Name	Watershed
349	Bull Creek Upstream of Tributary 7 (Franklin)	Bull Creek	5604	W Bouldin in West Bouldin Greenbelt	West Bouldin Creek
600	Bear Creek @ FM 1826 (South Fork)	Bear Creek	119	East Bouldin @ Elizabeth St.	East Bouldin Creek
612	Onion @ Driftwood (Hwy 150)	Onion Creek	3294	South Boggy @ Congress Ave	South Boggy Creek
48	Barton @ Hwy 71	Barton Creek	116	Shoal Creek @ 24 th St.	Shoal Creek
1114	Little Barton @ Great Divide Dr.	Little Barton Creek	3860	Little Walnut @ Georgian Dr.	Little Walnut Creek

Table 2. Stable and unstable site locations assessed in the Blackland Prairie ecoregion of Austin, TX.

Blackland Prairie Stable			Blackland Prairie Unstable		
Site No.	Site Name	Watershed	Site No.	Site Name	Watershed
1366	Onion @ SAR	Onion Creek	1210	Dry @ Wolf	Dry Creek East
1192	Gilleland @ 973	Gilleland Creek	888	Harris Branch @ Cameron Road	Harris Branch
5297	Wilbarger @ Brokenbrough-Mainstem	Wilbarger Creek	493	N. Boggy @ Delwau Lane	Boggy Creek
5298	Wilbarger @ Brokenbrough-Tributary	Wilbarger Creek	232	Marble @ Thaxton	Marble Creek
233	Rinard @ Bradshaw Road	Rinard Creek	503*	Walnut upstream of Freescale	Walnut Creek
-	-	-	898	Fort Branch @ Single Shot Circle	Fort Branch

* discarded from analysis due to its location in the transition between two ecoregions (see Results).

Parameter selection and collection

Data collection began in August 2012, when the majority of data was collected at each site. Additional data for the calculation of the entrenchment ratio was collected in January 2015. Fifteen geomorphic parameters were evaluated for each site and collection methods are described in Table 3. Qualitative parameters modified from Pfankuch (1975) were evaluated based on 4 categories consisting of excellent, good, fair, and poor (Appendix A). Many parameters used in this study were derived from the WPD's long-term EII monitoring that incorporates elements of channel stability evaluation while others were selected from relevant channel form assessment literature (i.e. channel cross section, entrenchment ratio, D50, Table 3).

Data analysis

Three of the fifteen assessed parameters consisted of multiple measurements (e.g. percentages of various substrate sizes and materials). Those parameters included bank material, pebble count, and bank angle. In order to optimize model inputs, these measurements were consolidated for analysis into a single representative parameter using various approaches described here.

The 10 bank material sub parameters were consolidated using a weighted distance measure formula (Equation 1) where *BM* is bank material, *Be* is bedrock; *Bo* is boulder; *Co* is cobble; *Gr* is gravel; *SSC* is sand, silt, and clay. *Rb* and *Lb* are prefixes for right and left bank designations, respectively. Briefly, the BM is low at sites with an even distribution of bank material and will increase as categories of a particular type of bank material increase in percentage. For example,

the BM at a site comprised mostly of boulders and bedrock will be slightly higher than a site with an even distribution of bank material, but the BM at a site comprised mostly of sand/silt/clay or gravel will be much higher than both other sites.

Equation 1:

$$BM = \sqrt{(RbBe^2 + LbBe^2) + (2 \times (RbBo^2 + LbBo^2)) + (3 \times (RbCo^2 + LbCo^2)) + (4 \times (RbGr^2 + LbGr^2)) + (5 \times (RbSSC^2 + LbSSC^2))}$$

Pebble counts for each site were consolidated by using the median particle size (D_{50}), calculated using Equation 2 where S^+ is the particle size at the top of the bin size, S^- is the particle size at the bottom of the bin size, P is the desired distribution percentage (i.e. 50%), P^+ is the percent of particles smaller than S^+ , and P^- is the percent of particles smaller than S^- (ODNR 2015). D_{50} is considered to be a conservative choice for the controlling particle for bed stability (Hawley and Vietz 2016 in review).

Equation 2:

$$D_{50} = 2^{\log_2(S^-) + [P - P^-] \times \frac{[\log_2(S^+) - \log_2(S^-)]}{[P^+ - P^-]}}$$

As the bank angle was measured on both sides of the stream at each location, the maximum bank angle measurement was chosen to represent the bank angle of each site. It was determined that using the most shallow bank angle would underestimate the stability of a site and an average of right and left bank angles may cause stable and unstable channels to have similar values. Thus, the steepest angle was selected as a conservative estimate to indicate channel access to the floodplain.

Table 3. Fifteen proposed geomorphic parameters and methods that were evaluated for each study site.

Functional Metric	Purpose and Method
Bank angle	Bank angle is indicative of the relative stability of the bank. Higher bank angles tend to be less stable than more gradual ones (Rosgen 1996). The bank angle of both the right and left bankfull banks was taken at single locations within a riffle representative of each site. Angles were measured using a clinometer or an electronic level.
Bank material	The size of materials comprising channel banks has been correlated with bank stability, with larger materials providing higher stability (Rosgen 1996). Visual percent estimates of 5 bank material categories (bedrock, boulders, cobble, gravel, and sand/silt/clay) for left and right bank were recorded for a total of 10 subparameters.
Pebble count	Dominant particle size and the size distribution of sediments within a channel are indicative of the amount of stream power within the system (Rosgen 1996). A Wolman pebble count was conducted within a riffle representative of the reach. One hundred random samples were taken and sized using a gravelometer. Counts for each site were then converted to their D ₅₀ value.
Channel type	Channel type was visually assessed based on three categories consisting of alluvial, bedrock, or structure-controlled channels. These factors affect the ability of the channel to adjust to changes in flow regime (Rosgen 1996).
Channel cross section	Repeated collection of cross-sectional areas at benchmarked locations allows for comparisons of changes in channel area over time, as well as inferences of rate of geomorphic changes (Rosgen 1996). Cross sections were measured across riffles that were indicative of the reach and uninfluenced by outliers or manmade features. Measurements were taken at stream bank grade changes using a measuring tape, stadia rod, and surveyor's level. Cross section measurements were not included in the data analysis.
Width/depth ratio	The width to depth ratio illustrates the distribution of energy within a channel and the ability of discharges to transport sediment (Rosgen 1996). Width to depth ratios were calculated using a bankfull width and its thalweg depth measured at one location within a riffle representative of the reach.
Entrenchment ratio (ER)	The entrenchment ratio (ER) is the ratio of the flood prone area to the surface width of the bankfull channel. Flood prone area was determined by measuring the width of the channel at twice the thalweg depth at bankfull. The ER indicates the extent to which the stream is vertically contained, or the degree it is incised into the valley floor (Kellerhalls et al. 1972). The ER also indicates how well the floodplain is connected to the stream (Stacey 2006). Measurements were taken at one location within a riffle representative of the reach.
Mass wasting	A visual qualitative analysis was conducted of the degree of existing/potential slope or mass movement of channel banks in the stream reach, modified from Pfankuch's channel stability evaluation (1975).

Table 3 continued.

Functional Metric	Purpose and Method
Bank protection from vegetation	A visual qualitative analysis was conducted of the amount and variety of vegetation covering channel banks in the stream reach, modified from Pfankuch's channel stability evaluation (1975).
Prevalence of obstructions, deflectors, and sediment traps	A visual qualitative analysis was conducted of the degree of embedded rocks and woody debris influencing a stable riffle/pool pattern, modified from Pfankuch's channel stability evaluation (1975).
Bank cutting	A visual qualitative analysis was conducted of the amount and height of bank cutting of channel banks in the stream reach, modified from Pfankuch's channel stability evaluation (1975).
In-channel deposition	A visual qualitative analysis was conducted of the amount of sediment deposition in a stream reach modified from Pfankuch's channel stability evaluation (1975).
Scouring and deposition	A visual qualitative analysis was conducted of the extent of stream bed scouring and bed material mobilization within the stream reach, modified from Pfankuch's channel stability evaluation (1975).
Stream reach slope	Slope is the change in elevation divided by the length of the stream reach. Slope affects the amount of potential energy in a stream reach (Rosgen 1996). Stream bed slope was estimated using ArcGIS to determine starting and ending elevation, and length of the reach. Length of reach was calculated using twenty times the bankfull width of the cross section.
Consolidation of particles	Packing of particles is an indicator of erosive potential in stream channels. A qualitative analysis of particle packing using a shuffle test was conducted, modified using Pfankuch's channel stability evaluation (1975).

Parameters were used as input into a random forest model of 5,000 trees, with the exception of channel type and channel cross section, to determine the importance of a parameter when classifying a location as stable or unstable. Channel type was all alluvial in the Blackland Prairie while a mix of alluvial and bedrock in the Edwards Plateau. This parameter was more of a descriptor of the ecoregion than an indicator for channel instability. The channel cross section, or area, was meant to be measured at two time periods in order to determine a change in the cross section area over time. No second channel cross section was measured. Thus, channel cross section measurement was not included in the current development of a stability index, but can be used to assess channel evolution in these reaches over time.

Random forest analysis forms multiple classification trees on a data set which can be used to identify important predictor variables in regards to a response variable and accurately classify response variables into categories. 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 OOB 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 (Breiman 2001). To measure variable importance, OOB data is used to randomly permute values of each parameter while other parameters are held constant. Variable importance is then calculated as the difference between the OOB error estimate when the parameter is permuted and the OOB error estimate from the original model (Breiman 2002).

Important parameters were used to develop the multi-metric Stream Stability Index for each ecoregion. 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, a criticism to using multi-metric indices as a measure of environmental health is the indices can lack interpretability or usefulness (Suter 1993, Ohe et al. 2007). Bayesian decision theory was used to calculate the probability that a site would be classified as stable based on the collected measurements of important parameters. This probability was calculated for each parameter at each site. The Stream Stability Index (SSI) was computed as a weighted sum of each probability.

In general, the probability of a site to belong to a classification c given a set of environmental metrics x can be denoted as $P(c|\mathbf{x})$. Bayes rule is then used to solve for this probability (Equation 3):

Equation 3:

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

In the above formula $P(c)$ is the prior probability of a site to be in a specific classification, $P(\mathbf{x}|c)$ is the likelihood of observing the vector \mathbf{x} given that the site is in a specific classification, and $P(\mathbf{x})$ is the probability of observing the vector \mathbf{x} . More specifically, the calculation of the SSI consists of two classifications (stable, Equation 4, and unstable, Equation 5) and uses the naïve Bayes classifier to solve for $P(\mathbf{x})$ and $P(\mathbf{x}|c)$ under the assumption that prior probabilities will be uninformative or $P(c)$ is 0.5:

Equation 4:

$$P(c_{\text{stable}}|\mathbf{x}) = \frac{\prod_{j=1}^N P(x_j|c_{\text{stable}})}{\prod_{j=1}^N P(x_j|c_{\text{stable}}) + \prod_{j=1}^N P(x_j|c_{\text{unstable}})}$$

Equation 5:

$$P(c_{\text{unstable}}|\mathbf{x}) = \frac{\prod_{j=1}^N P(x_j|c_{\text{unstable}})}{\prod_{j=1}^N P(x_j|c_{\text{stable}}) + \prod_{j=1}^N P(x_j|c_{\text{unstable}})}$$

where x is a matrix of N independent environmental metrics. $P(c_{\text{stable}}|\mathbf{x})$ is the probability of a site to belong to the group of reference sites given the collected environmental metrics at the site. Because of the qualitative characteristics of the metrics collected for this study the $P(c_{\text{stable}}|\mathbf{x})$ was either 0 or 1 for every site when all metrics were combined. While this is exactly what is desired from a classification tool, it did not provide a range of probability for each site to be stable. Thus $P(c_{\text{stable}}|\mathbf{x})$ was computed for each parameter individually at every site and weighted based on the decrease in accuracy computed from the random forest model. The Stream Stability Index was the sum of these weighted probabilities. Index scores greater than 50 were categorically classified as stable while values less than 50 were classified as unstable.

Results

The site on Walnut Creek at Freescale is located along the edge of the Edwards Plateau and Blackland Prairie and thus portrays characteristics of both ecoregions. During initial stages of analysis, it was noted that a separate random forest model for each ecoregion was necessary to accurately classify locations as stable or unstable. Walnut Creek upstream of Freescale was dropped from further analysis because environmental parameters collected at the site did not conform to either classification most likely due to its location in the transition zone between the two ecoregions.

Results of the model analysis varied for each ecoregion. The initial error rate for the model using all parameters within the Edwards Plateau was 10% while the error rate for a similar model within the Blackland Prairie was 20%. In addition, parameters varied in importance between the Edwards Plateau and Blackland Prairie regions (Figure 2). For example, permutation of the

Obstructions and Deflectors parameter in the Edwards Plateau model caused the second biggest difference in model error (was the second most important parameter) but was not important in the Blackland Prairie model. Scouring and Deposition and Mass Wasting were the only parameters that were within the top 5 important parameters for both Edwards Plateau and Blackland Prairie models. Individual figures for the top ten parameters of each ecoregion, excluding Scouring and Deposition and D_{50} are located in Appendix B.

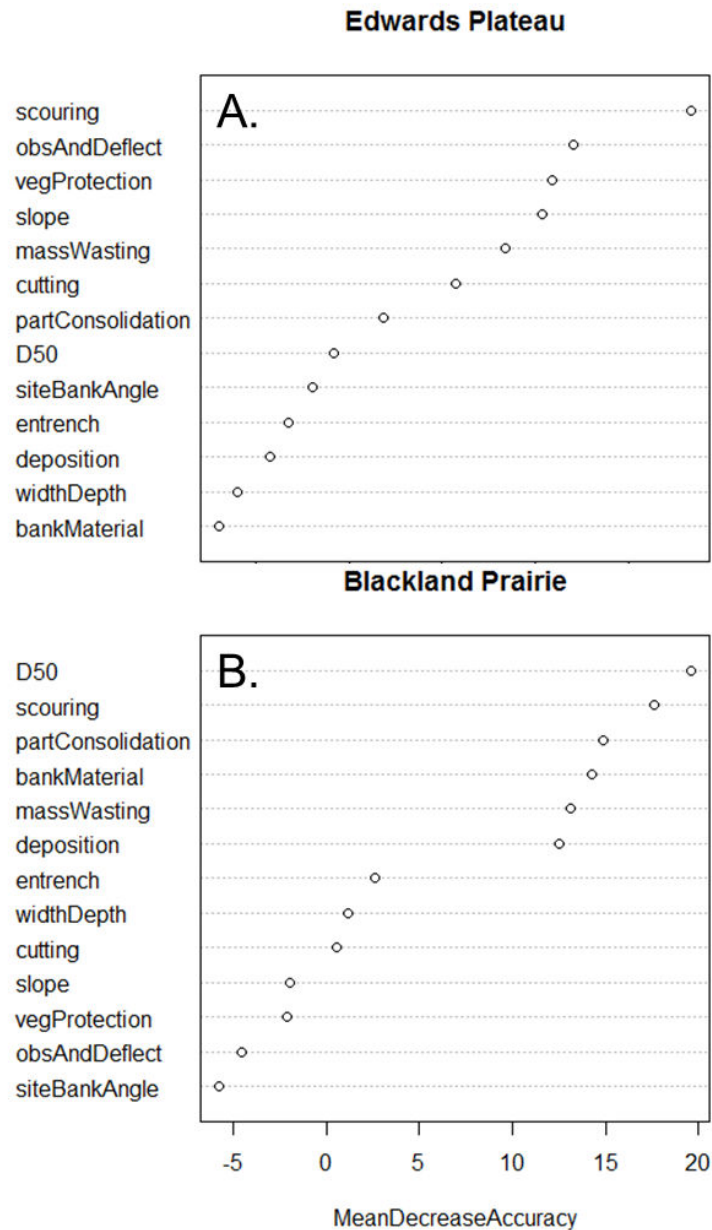


Figure 2. Mean decrease in accuracy of the “stable” and “unstable” site classification model, or change in OOB error estimate, when a particular parameter (Y-axis) was excluded from the (A) Edwards Plateau and (B) Blackland Prairie models. Large change in OOB error estimates results in a mean decrease in model prediction accuracy and implies importance of that parameter, which are listed in order, top being most important, bottom least.

Overall, qualitative parameters better improved model accuracy when compared to the measured quantitative parameters, with the exception of the D_{50} in the Blackland Prairie. In the Edwards Plateau, four of the top five predictive parameters were qualitative modified Pfankuch (1975) categories while three of the top five predictive parameters were qualitative in the Blackland Prairie.

For both ecoregions, when the Scouring and Deposition visual assessment parameter was inserted into a random forest model with no other inputs the classification error rate was 0%. In contrast, the In-channel Deposition parameter has a 90% error rate when placed in both models as an individual parameter. For the Blackland Prairie ecoregion, a similar model set up using the D_{50} resulted in a classification error rate of 0% as well. In other words, the Scouring and Deposition parameter was able to correctly classify sites as stable or unstable with 100% accuracy in both Edwards Plateau and Blackland Prairie ecoregions and the D_{50} parameter was able to do the same in the Blackland Prairie. Stable sites for both ecoregions were categorized as either excellent or good using the Scouring and Deposition parameter while unstable sites were classified as fair or poor (Figure 3, Appendix A). The D_{50} calculated for stable sites within the Blackland Prairie ranged from 21.83 mm to 51.51 mm while the unstable sites had values that ranged from 0.06 mm to 13.76 mm (Figure 4).

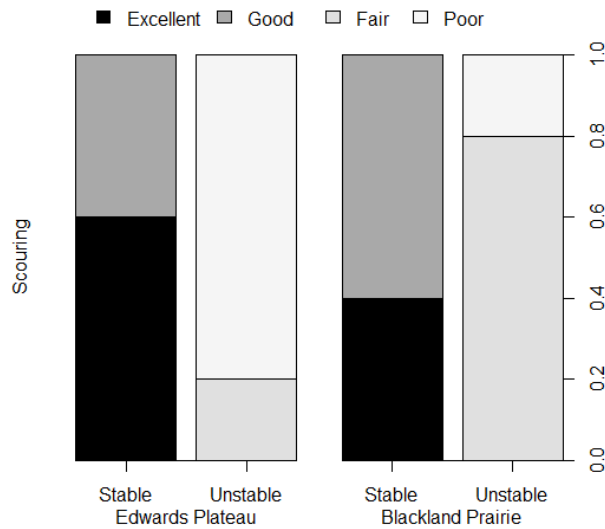


Figure 3. Spine plot of Scouring and Deposition for stable and unstable stream reach sites in the Edwards Plateau and Blackland Prairie ecoregions of Austin, TX. Decimal values represent the proportion of sites within a category with a given value for scouring (0.6 of stable sites in the Edwards Plateau were ranked Excellent).

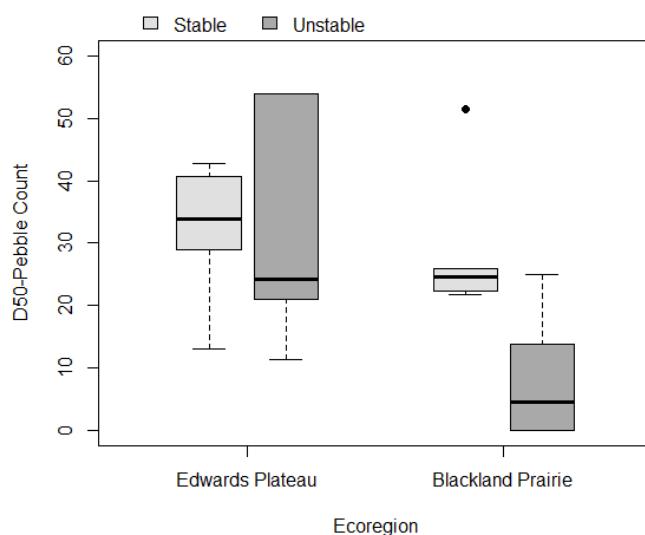


Figure 4. Box plot of D50 for pebble counts collected for stable and unstable stream reach sites in the Edwards Plateau and Blackland Prairie ecoregions of Austin, TX. Box-and-whisker plots delineate the mean (dark line), interquartile range (shaded area), 5-95% range (whiskers), and outliers (black dots).

Discussion

The importance of parameters within each ecoregion is likely related to the geography of the area. The geography of the Edwards Plateau consists of a steep rolling topography comprised of shallow soils resting on hard limestone. Large bank and channel materials were common in the Edwards Plateau and the Obstructions and Deflectors parameter was likely easier to visually assess in this ecoregion due to a more noticeable contrast of the presence and absence of large rocks at each site. The topography also yielded more noticeable differences between the stream reach slopes of stable and unstable sites compared to Blackland Prairie sites (Appendix B, Figure B6(D)).

In contrast, the geography of the Blackland Prairie is comprised of a generally flat topography consisting of chalk (soft limestone), clay, sand, or alluvium. The deep soils cause the channels to be driven by alluvial processes, which is likely the reason parameters associated with bank/channel substrate and erosive potential accurately classified sites (e.g. D₅₀, Bank Material, Consolidation of Particles, and Mass Wasting).

Regardless of geographic differences, Bank Protection from Vegetation was expected to be an important parameter for classification in both ecoregions. The parameter did not perform well in the Blackland Prairie model. One reason for the lack of importance was that the site in Gilleland Creek at FM 973, a site classified as stable, was given a “poor” rating. Sapling and shrub vegetation was indeed sparse on the banks of the site; however, armoring of the banks was provided by extensive root networks of mature trees near the banks, which were not accounted for during the site assessment (Figure 5). Canopy cover from the abundance of mature trees at the site may also prevent growth of a strong ground cover and understory community, resulting in a poor score for that parameter in an otherwise very stable bank.



Figure 5. Photo of the right bank of site 1192, Gilleland Creek at FM 973. Vegetation was sparse along the bank, but dense root networks of mature trees provided armoring.

Qualitative parameters, particularly those modified from Pfankuch (1975), were more successful at classifying stable and unstable sites than most quantitative parameters (i.e. width to depth ratio, entrenchment ratio, and bank angle). One reason for this may be that the quantitative measures were not representative of the channel morphology. These three parameters were measured at a single location within each reach while qualitative parameters were integrated over the entire reach. Bank Angle may have performed poorly because only the steepest angle for each site was used in the models. While using the steepest angle was a more conservative option than the shallow angle or an average of right and left bank angles, it may have caused channel access to the floodplain to be underestimated yet we determined the conservative estimation was the best option for the analysis. In addition, a single categorical assessment can act as a surrogate for other unmeasured parameters. For example, the Scouring and Deposition parameter probably includes geomorphic and hydrologic influences such as channel incision, peak flow rate, particle consolidation, and bedload particle size, which may be the reason for its universal accuracy in this study.

Evidence of scouring can also indicate reductions in ecosystem function. Removal and/or movement of bedload sediment can decrease hyporheic exchange impacting nutrient retention, temperature modulation and metabolism, and impact the capacity for pollutant synthesis as well

as diminish foraging and refuge for macroinvertebrates and fish (Richter 2011, Vietz et al. 2014). Hence, the success and utility of the Scouring and Deposition parameter should strongly suggest continued data collection at EII sites in order to complement WPD biological monitoring data.

There was a large contrast in site classification accuracy when using Scouring and Deposition as opposed to In-channel Deposition. Both parameters are designed to evaluate deposition, yet the classification error rates of each were different in single variable models. Categorical descriptions of excellent, good, fair, and poor for Scouring and Deposition may be more easily interpreted and evaluated by the observer than ones given for In-channel Deposition, particularly if the observer has no historical knowledge of the site. For example, the categories for Scouring and Deposition are described as the percentage of the stream reach that is affected by scour and deposition (Appendix A). In contrast, the wording for In-channel Deposition categories such as “new increase in bar formation” or “moderate deposition of new gravel and coarse sand on old/new bars” suggest that some historical site knowledge may be required for accurate evaluation (Appendix A). The accuracy of In-channel deposition could increase with repeated assessment and photo documentation of each site, but baseline evaluations with limited historical documentation, such as those conducted in this study, could be less accurate.

Cross section measurements were taken at each site for reference, but not included in the random forest analysis. This is because multiple cross section measurements are required to compute shifts in the cross section before the measurements are useful in predicting channel stability. However, a comprehensive report on long-term cross section measurements in Austin streams is currently being drafted by the WPD Stormwater Treatment and Stream Restoration (STSR) section and the results from the cross section assessments should be compared to the findings from this study.

Stream Stability Index Development

Based on results from the models of both ecoregions, the development of an ecoregion-specific Stream Stability Index for both the Edwards Plateau and Blackland Prairie is appropriate. Each ecoregion has important parameters that would confound the other ecoregion if included into a single index. For example, the D_{50} parameter perfectly classified stable and unstable sites in the Blackland Prairie but not in the Edwards Plateau. However, analysis suggests that one parameter, Scouring and Deposition, was highly successful at accurately classifying a stream as stable or unstable across both ecoregions. It appears that the Scouring and Deposition parameter may act as a surrogate for many other hydrologic, geomorphic, and ecological parameters, and be a good tool to assess stream instability impacts on aquatic life in Austin.

Challenges associated with qualitative visual assessments, like the Scouring and Deposition parameter, include user bias and over-generalization. These challenges may make consistent evaluations over the long-term difficult unless consistent training is conducted prior to evaluation. In the case of this variable, a variety of staff have used it at all EII sites during multiple visits over the past 15 years, providing an excellent data set for assessing variability. Throughout the data set, Scouring and Deposition has been observed to vary within two adjoining categories (Excellent-Good, Good-Fair, Fair-Poor) at a given site location. According to this analysis, most sites measured over the past 15 years could be classified as either stable or

unstable with a few sites crossing between stable and unstable throughout the time period (Good-Fair sites).

Although most evaluated parameters had some range of error for classifying stable and unstable sites, the inclusion of multiple top performing parameters into an index would likely increase the robustness of the index when applied to a larger scale. Thus, based on our analysis, we recommend parameters listed in Table 4 be included in Stream Stability Indices specific to each ecoregion. All study sites were correctly classified using ecoregion specific models made up of the suggested parameters.

Table 4. Suggested parameters for ecoregion specific Stream Stability Indices. Models utilizing these specific parameter sets were able to correctly classify all stable and unstable sites in their respective ecoregions.

Suggested Edwards Plateau Stream Stability Index Parameters	Suggested Blackland Prairie Stream Stability Index Parameters
D ₅₀	D ₅₀
Scouring and deposition	Scouring and deposition
Consolidation of particles	Consolidation of particles
Mass wasting	Mass wasting
Bank protection from vegetation	Bank material
Prevalence of obstructions, deflectors, and sediment traps	-

SSI scoring results for the Edwards Plateau and Blackland Prairie regions can be seen in Table 5 and 6, respectively. Scores range from 0 to 100 and reflect the sum of the weighted individual parameter probabilities based on the decrease in accuracy computed from the random forest model. Index values greater than 50 were categorically classified as stable while values less than 50 were classified as unstable. Observer error on one of the collected parameters will not cause misclassification of a site so the scoring should be robust. However, the scoring system should be validated at sites that were not included in this study.

Table 5. Classification scoring results of the Edwards Plateau Stream Stability Index. The site type category indicates initial stability category.

Site No	Site Name	Site Type	Score
48	Barton Creek @ Hwy 71 Downstream of Little Barton	Stable	88
116	Shoal Creek @ 24th Street	Unstable	9
119	East Bouldin Creek @ Elizabeth St	Unstable	22
349	Bull Creek Upstream of Tributary 7 (Franklin)	Stable	82
600	Bear Creek @ FM 1826 (South Fork)	Stable	83
612	Onion Creek near Driftwood (Hwy 150)	Stable	76
1114	Little Barton Creek @ Great Divide Dr	Stable	81
3294	South Boggy @ Congress Ave	Unstable	15
3860	Little Walnut Creek @ Georgian Dr	Unstable	15
5604	West Bouldin Creek in West Bouldin Greenbelt	Unstable	35

Table 6. Classification scoring results of the Blackland Prairie Stream Stability Index. The site type category indicates initial stability category.

Site No	Site Name	Site Type	Score
232	Marble Creek @ Thaxton Road (M2)	Unstable	37
233	Rinard Creek @ Bradshaw Road	Stable	72
493	North Boggy Creek @ Delwau Lane	Unstable	13
503	Walnut Creek Upstream of Freescale	Unstable	43
888	Harris Branch Creek @ Cameron Road	Unstable	16
898	Fort Branch Creek @ Single Shot Circle	Unstable	25
1192	Gilleland Creek @ FM 973	Stable	81
1210	Dry Creek @ Wolf Lane	Unstable	16
1366	Onion Creek @ South Austin Regional WWTP	Stable	81
5297	Wilbarger Bottomlands	Stable	73
5298	Unnamed Tributary to Wilbarger Headwaters	Stable	67

Time and effort involved using the suggested indices to assess a stream site is estimated to range from 1 to 1.5 hours. However, if additional geomorphic and hydrologic data is collected (e.g. channel cross section measurements and stream slope) then site assessment time could extend to 3 hours.

Recommendations

- Each ecoregion specific model should be validated by assessing additional known stable and unstable sites historically surveyed by WPD STSR section.
- Compare the developed Stream Stability Indices with WPD Rapid Geomorphic Assessment (DEC 2015) scores at applicable site locations to determine the variation between these two qualitative methods.
- Determine $Q_{critical}$ flow for EII sites using guidelines suggested in Hawley and Vietz (2016 in review). This will require a D_{50} , channel slope, boundary shear stress, and hydraulic radius, which are data that may be partially derived from models and/or collected in the field. This quantitative approach could help watershed managers gain a better understanding of how to create a regionalized stormwater management strategy that more closely mimics the hydrograph of Austin streams prior to urban development.
- Continue to collect channel cross section, flood prone area, and bankfull width measurements at sites during SSI evaluations, but do not directly include this data in the actual index. These measurements will allow width to depth and entrenchment ratios as well as the hydraulic radius to be calculated as supplemental data, which would be beneficial data for the WPD STSR section. In addition, it is recommended that the aforementioned parameters along with bank angle be based on an average of at least three points in a reach rather than one in order to more accurately reflect the morphology of the channel.
- Devise an abbreviated method for evaluating stream reach slope. Using tools such as a handheld elevation GPS to delineate the thalweg profile of a stream reach may be more accurate than the GIS-based method used in this study, particularly if digital elevation models are not current. Even though stream reach slope was not a top performing parameter in this study, WPD considers the parameter important for calculating $Q_{critical}$. Hawley and Vietz (2016 in review) recommend that thalweg profiles should, at a minimum, extend to the crests of the nearest up and downstream riffles.
- It is recommended that a method to estimate Manning's N, channel roughness coefficient, in and around the main site riffle be adopted. Estimating stream flow using Manning's equation is one other component that is important for calculating $Q_{critical}$. Tables for estimated N values related to different channel types can be found in Chow (1959).
- If the Bank Protection from Vegetation parameter is to be utilized for assessment in the future, the method should include clear guidance on how to assess deeply shaded areas with little to no ground cover but robust armoring by root masses.

Acknowledgements

We would like to thank Nathan Leber for his contributions during the initial study design and data collection phases of this project.

References

- Breiman, L. 2001. Random forests. *Machine Learning*, 45(1): pp 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.
- City of Austin (COA). 2002. Environmental Integrity Index Water Quality Technical Assessment Methodology. City of Austin, Watershed Protection Department, Environmental Resources Management Division, Water Resource Evaluation Section: CM-02-01.
- Chin, A. 2006. Urban transformation of river landscapes in a global context. *Geomorphology* 79(1) 460-487.
- Chow, Ven Te. 1959. *Open-channel hydraulics*. McGraw-Hill Book Company, Inc., New York, NY, 680 pp.
- Department of Environmental Conservation (DEC). 2015. New York state stormwater management design manual. Appendix J: Geomorphic assessment.
- Hawley, R.J. and Vietz, G. J. 2016. Addressing the urban stream disturbance regime. *Freshwater Science* (in review).
- Hooke, J. M. 1986. Applicable and applied geomorphology of rivers. *Geography* 71(1) 1-13. Geographical Association.
- Karr, JR and EW Chu. 1997. *Biological Monitoring and Assessment: Using Multimetric Indexes Effectively*. EPA 235-R97-001. University of Washington, Seattle.
- Kellerhals, R., Neill, C. R., and Bray, D. I. 1972. Hydraulic and geomorphic characteristics of rivers in Alberta. Research Council of Alberta, River Engineering and Surface Hydrology Report 72-1.
- Knighton, D. 1998. *Fluvial forms and processes: A new perspective*. London, UK. Hodder Education. 383 p.
- Knox, J. C. 2006. Floodplain sedimentation in the Upper Mississippi Valley: Natural versus human accelerated. *Geomorphology* 79 (1) 286-310.
- Ohe, P. von der, A. Prüß, and R. Schäfer. 2007. Water quality indices across Europe—a comparison of the good ecological status of five river basins. *Journal of environmental monitoring* 9(9) 970-978.

- Ohio Department of Natural Resources (ODNR), Division of Soil and Water Resources. 2015. Calculating D50. http://soilandwater.ohiodnr.gov/portals/soilwater/data/doc/Calculating_D50.doc.
- Omernik, J. M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* 77(1) 118-125.
- Pfankuch, D. J. 1975. Stream reach inventory and channel stability evaluation. USDA Forest Service, R1-75-002. Government Printing Office#696-260/200, Washington D.C.
- Richter, A. 2011. Linking biologic metrics to hydrologic characteristics in Austin, Texas streams. City of Austin, Watershed Protection Department, Environmental Resource Management, SR-11-15.
- 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.
- Rosgen, D. 1996. Applied river morphology. Wildland Hydrology, Pagosa Springs, CO.
- Stacey, P. B., Jones, A. Catlin, J. C., Duff, D. A., Stevens, L. E., and Gourley, C. 2006. User's guide for the rapid assessment of the functional condition of stream-riparian ecosystems in the American Southwest. Wild Utah Project.
- Suter, GW. 1993. A critique of ecosystem health concepts and indexes. *Environmental Toxicology and Chemistry* 12: 1533–1539.
- Vietz, G. J, Sammonds, M. J., Walsh, C. J., Fletcher, T. D., Rutherford, I. D., Stewardson, M. J. 2014. Ecologically relevant geomorphic attributes of streams are impaired by even low levels of watershed effective imperviousness. *Geomorphology* 206 (1) 67-78.

Appendix A

Table A7. Qualitative geomorphic parameters adapted from Pfankuch (1975) that were used in the study. Parameters were placed into 1 of 4 categories consisting of excellent, good, fair, or poor.

	Excellent	Good	Fair	Poor
Mass wasting (existing or potential)	No evidence of past or potential for mass wasting into channel	Infrequent or very small, mostly healed over, low future potential	Moderate, frequency and size, with some raw spots	Frequent or large, causing sediment loss nearly year long
Bank protection from vegetation	90% plant density, vigor and variety suggest deep, dense root mass	70 - 90% density, fewer species or lower vigor	50 - 70% density, lower vigor and fewer species, shallow/discontinuous root mass	< 50% density, fewer species, less vigor, poor, discontinuous, shallow root mass
Obstructions, deflectors, sediment traps	Rocks, old logs firmly embedded, flow pattern of riffle/pool stable w/o cutting or deposition	Moderately frequent, minor pool filling, obstructions and deflectors newer and less firm	Less frequent and unstable obstructions and deflectors, move with high water causing bank cutting and filling of pools	Free of obstructions and deflectors, cause bank erosion yearlong, sediment traps full, channel migration occurring
Cutting	Little to none evident, infrequent raw banks less than 6" high	Some, intermittent outcurves and constrictions, raw banks up to 12" high	Significant, cuts 12-24" high, root mat overhangs and sloughing evident	Almost continuous cuts, some over 24" high, failure of overhangs frequent
In-channel deposition	Little or no enlargement of channel or point bars	Some new increase in bar formation, most from coarse gravels	Moderate deposition of new gravel and coarse sand on old/new bars	Extensive deposition of predominantly fine particles, accelerated bar development
Consolidation or particle packing: shuffle test	Assorted sizes tightly packed and/or overlapping; little to no movement of substrate with test	Moderately packed with some overlapping; low to moderate movement of substrate	Mostly a loose assortment with no apparent overlap; moderately high movement of substrate	No packing evident, loose assortment, easily moved; high movement of substrate with little resistance
Scouring and deposition	Less than 5% of bottom affected by scouring and deposition	5-30% affected, scour at constrictions and where grades steepen, some deposition in pools	30-50% affected, deposition and scour at obstructions, constrictions, and bends, pools filling	More than 50% of bottom is in a state of flux or change nearly yearlong

Appendix B

The following section contains figures of the top 10 predictive parameters for both Edwards Plateau and Blackland Prairie models excluding scouring and deposition and D_{50} (see Figure 3 and 4). Spine plots were generated for categorical variables while box-and-whisker plots were created for continuous variables.

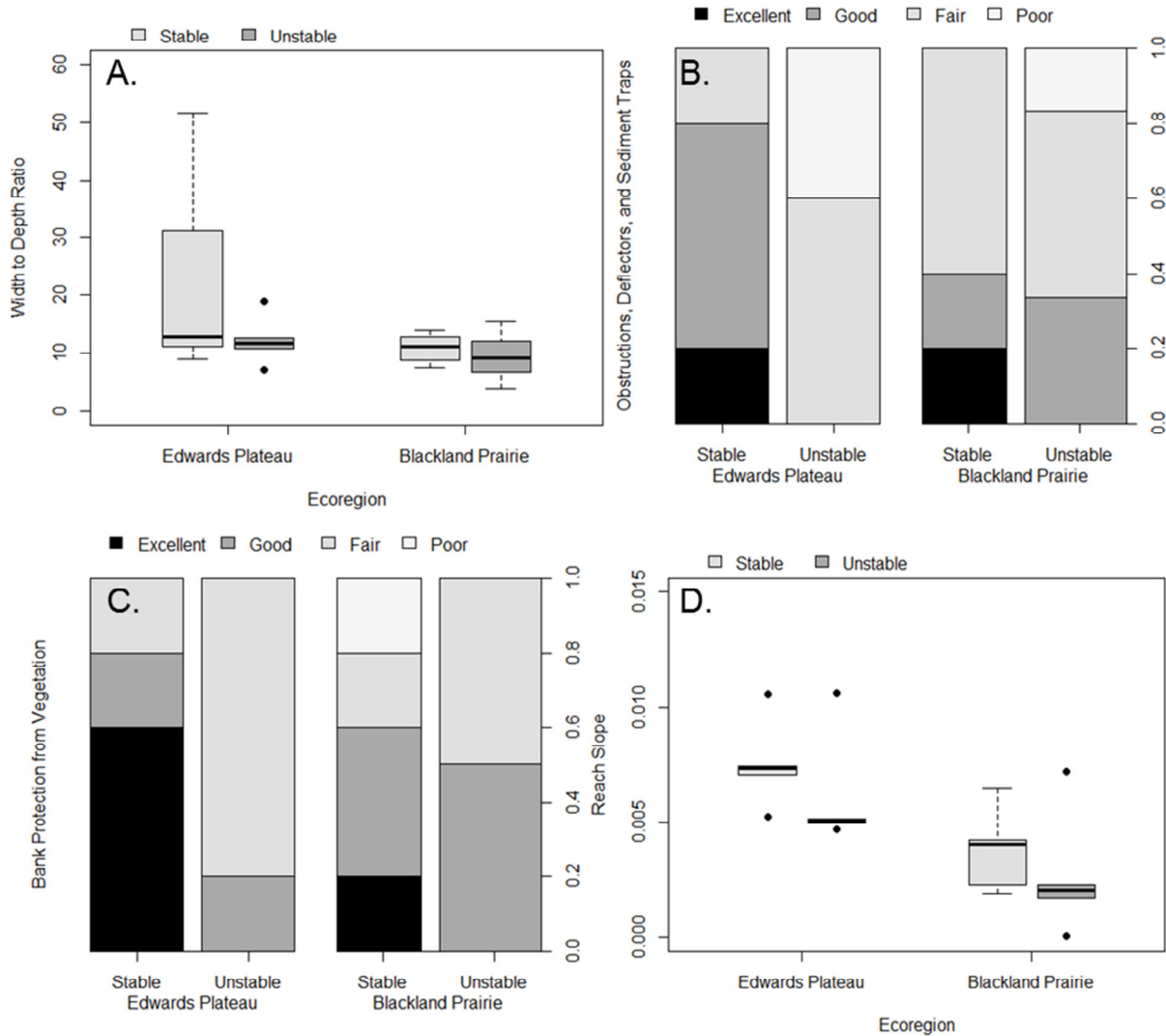


Figure B6. Width to depth ratio (A), obstructions, deflectors, and sediment traps (B), bank protection from vegetation (C), and stream reach slope (D) parameters for stable and unstable stream reach sites in the Edwards Plateau and Blackland Prairie ecoregions of Austin, TX. Decimal values are shown for reference on the right Y axes of spine plots. Box-and-whisker plots delineate the mean (dark line), interquartile range (shaded area), 5-95% range (whiskers), and outliers (black dots).

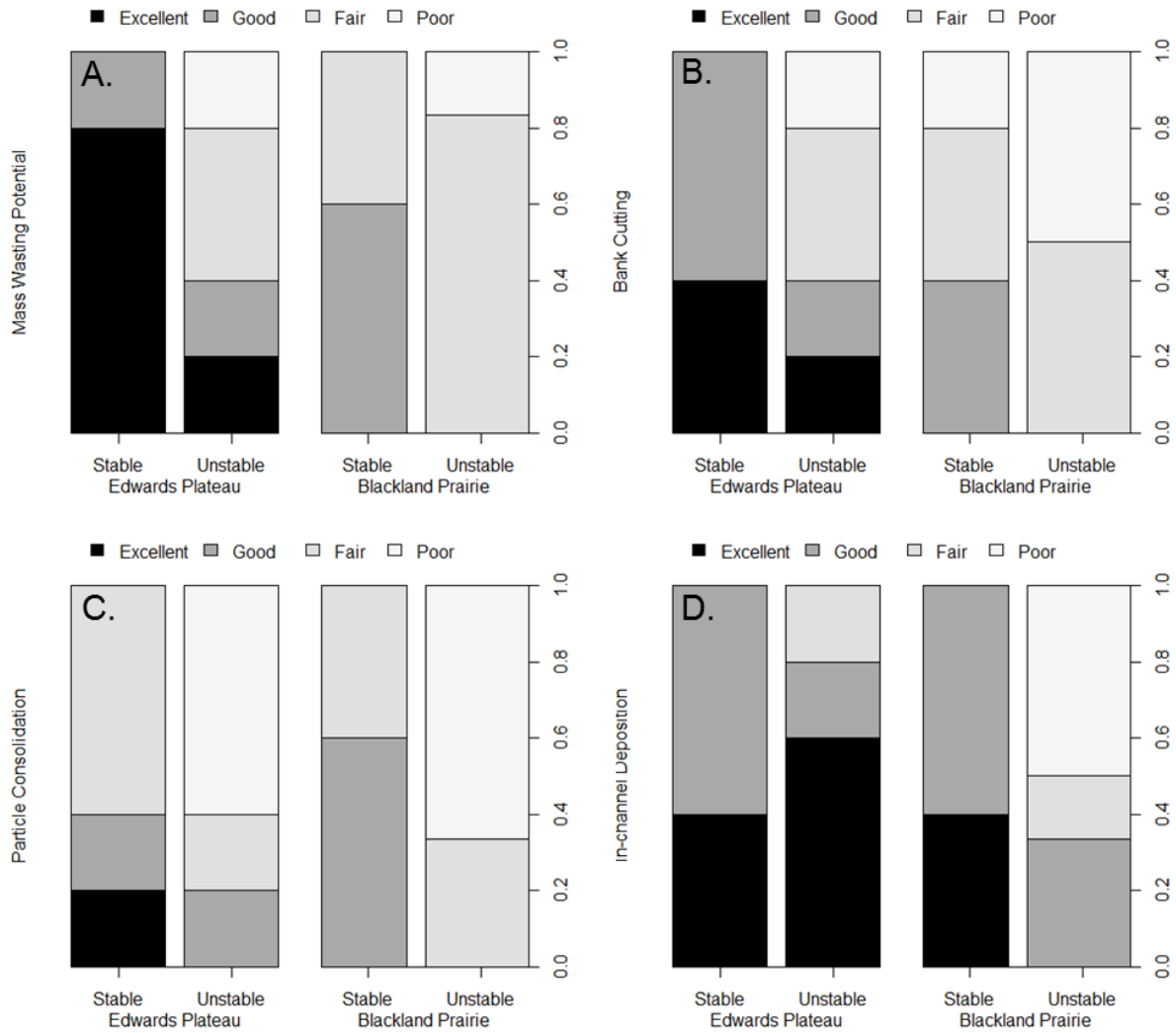


Figure B7. Mass wasting potential (A), bank cutting (B), particle consolidation (C), and in-channel deposition (D) parameters for stable and unstable stream reach sites in the Edwards Plateau and Blackland Prairie ecoregions of Austin, TX. Decimal values are shown for reference on the right Y axes of spine plots. Box-and-whisker plots delineate the mean (dark line), interquartile range (shaded area), 5-95% range (whiskers), and outliers (black dots).

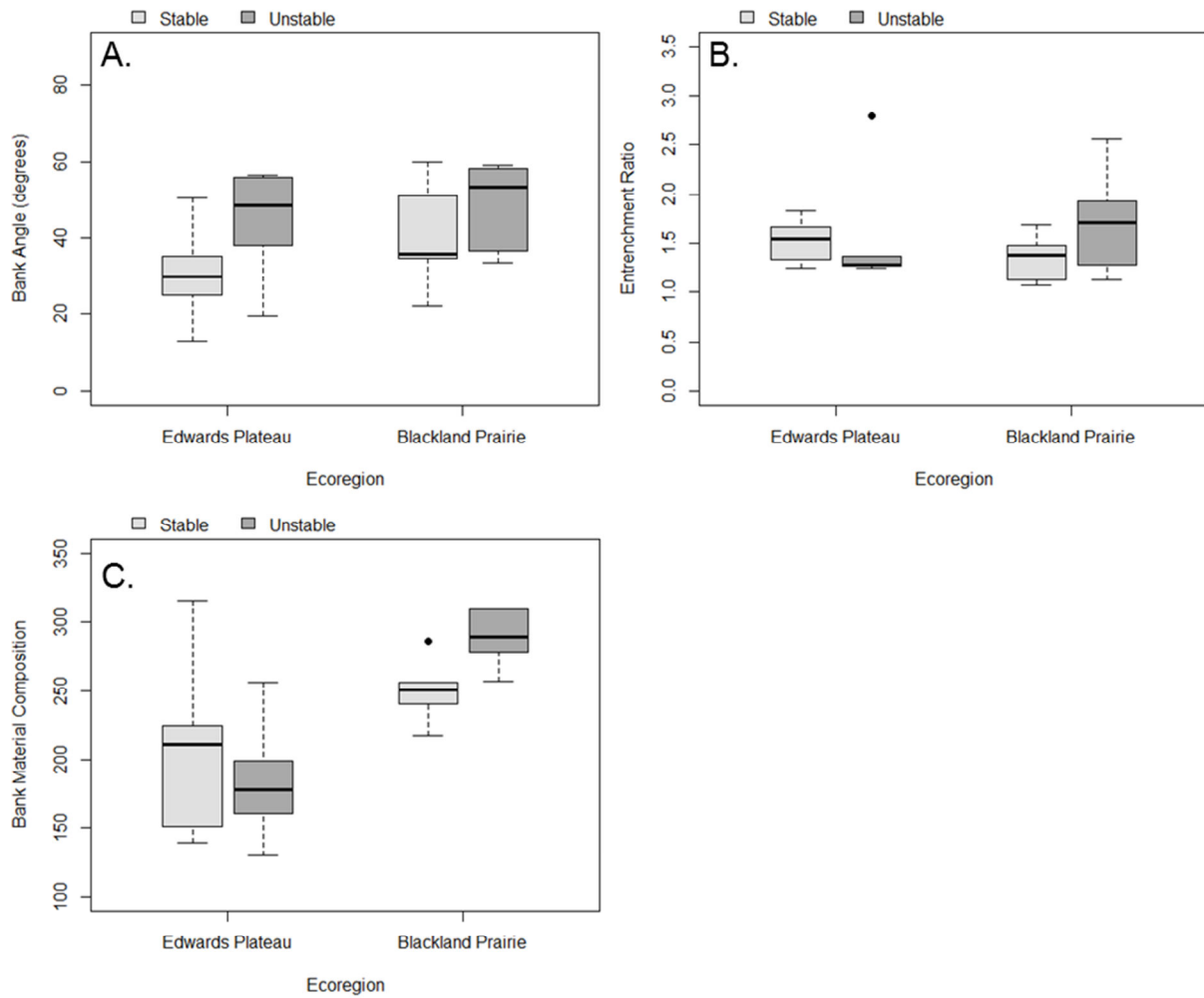


Figure B8. Bank angle (A), entrenchment ratio (B), and bank material composition (C) parameters for stable and unstable stream reach sites in the Edwards Plateau and Blackland Prairie ecoregions of Austin, TX. Decimal values are shown for reference on the right Y axes of spine plots. Box-and-whisker plots delineate the mean (dark line), interquartile range (shaded area), 5-95% range (whiskers), and outliers (black dots).