A Review of the Bottom Deposits Standard

Improving Evaluation of Perennial Streams

July 2018

By Catherine Gullo

Arizona Department of Environmental Quality
Open File Report
ADEQ Water Quality Division Surface Water Section
Monitoring Unit, Standards & Assessment Unit
1110 West Washington St.
Phoenix, Arizona 85007-2935

ADEQ
Arizona Department
of Environmental Quality

ACKNOWLEDGEMENTS

Funding for this study provided by the Arizona Department of Environmental Quality (ADEQ).

Thanks to all field personnel over the years who spent long hours collecting samples, including Aiko Condon, John “Trey” Dempsey, Bryant Dickens, Susan Fitch, Lee Johnson, Jason Jones, Sam Rector, Meghan Smart, Patti Spindler, and John Woods.

Report Review: Jason Jones, Patti Spindler.
Report Cover: http://www.krisweb.com/stream/sediment.htm

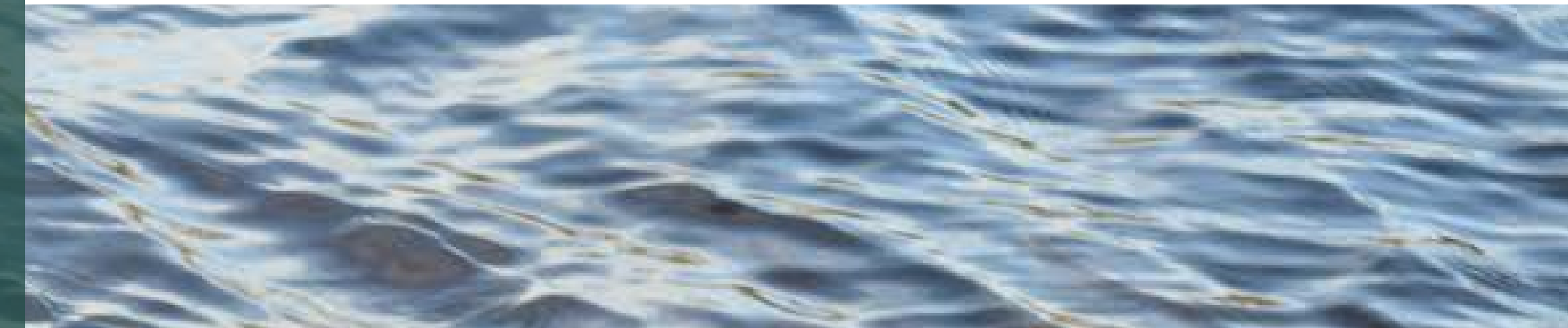

TABLE OF CONTENTS

INTRODUCTION

Excessive fine sediment has a detrimental effect on the health of in-stream ecosystems. Aquatic invertebrates and fish species make use of pools and interstitial spaces of larger particles, such as cobbles and gravel found in riffle habitats, for reproducing and early life stage development. Interstitial habitats are crucial zones for fish spawning, sediment surface attachment by larval stages of benthic macroinvertebrates, and oxygen transfer in micro- and macrobiotic communities. Fine sediment, defined here as particles less than 2 mm in diameter, can degrade and eliminate these important habitats by filling the spaces between larger sediment particles and excess sediment deposition in pools. Reference condition streams with low percent fine sediment in stream bottom substrate tend to have more abundant and diverse macroinvertebrates (Chapman and McLeod, 1987). Species of Heptageneid mayflies are sensitive to excess sediment and considered intolerant to highly embedded conditions; whereas species of Chironomus midges are more tolerant to high sediment levels (Relyea, et al., 2000).

Sedimentation was identified as a leading cause of water quality impairment in the country (USEPA, 2002). Major inputs of sediment into surface water are from anthropogenic sources, such as construction and mining. The Clean Water Act sets a national goal for water quality that protects both public health and aquatic and wildlife, through its objective "to restore and maintain the chemical, physical and biological integrity of the nation's waters." To address this objective and implement guidance to interpret the existing narrative bottom deposit standard (R-18-11-108), ADEQ developed bottom deposits numeric criteria for warm and cold water perennial streams in the 2009 Water Quality Standards, along with associated Implementation Procedures (Spindler, 2015b). These criteria were based on literature that illustrated the adverse effects of sediment in streams on the populations of benthic macroinvertebrates and fish. For instance, decline in aquatic insect species abundance and diversity was associated with habitat loss caused by sedimentation, and increases in fine sediment in spawning riffles related to decline in survival of salmonid embryos (Waters, 1995; Chapman and McLeod, 1987; Bjornn et. al., 1977). Criteria were established for warm and cold water streams in Arizona based on the results of these studies, in order to better protect aquatic life: sediment levels of ≤30% fines for cold water streams and ≤50% fines for warm water streams. These water quality standards are meant to protect the aquatic and wildlife (A&W) cold and warm designated uses, and are measured as a percentage of fine sediment present in the substrate of the wetted channel.

The current numeric criteria, developed to protect Arizona's aquatic life, are not based on empirical data gathered in the state. The current criteria may not be representative in depicting the specific conditions and stressors of stream systems of Arizona, since they are based on datasets from the Pacific Northwest and Idaho. The current numeric standards are suspected to be too high to make accurate impairment decisions, causing Type II errors; eg. Not finding impairment when the stream ecosystem may in fact be sediment stressed. This can directly affect impairment decisions made by the state, causing an oversight of stream reaches that would be considered exceeding the bottom deposits listing criteria. If a stream is not deemed impaired, the sediment issue cannot be monitored and remedied. The 2012 Statewide Assessment of Arizona Streams report provided preliminary reference site thresholds that suggested that the current bottom deposits standards thresholds may not be representative of true reference stream conditions in Arizona (Jones, 2012). That 2012 study suggested lower thresholds for fine sediment; cold water streams are considered in good condition if below 16% fines, and below 20% fines for warm water streams. These thresholds are approximately 50% less than current bottom deposits criteria.

The primary objective of this study was to evaluate the bottom deposit standards thresholds, using ADEQ datasets to determine whether the existing criteria need to be updated. Empirically derived datasets will better represent reference stream conditions in Arizona and provide improved protection of aquatic and wildlife designated uses.
We will analyze historic bottom deposits data in order to explore the relationships between stream sediment conditions and biological health, to ultimately develop new bottom deposit standards that better protect aquatic life in Arizona. The study will also assess the sampling methods employed by ADEQ, and make recommendations for improving data collection and analysis to provide a more holistic and representative dataset for continued development of new standards.

SAMPLING AND ANALYSIS METHODS

The ADEQ *Standard Operating Procedures (SOP) for Water Quality Sampling* (Jones, 2018) describes the sample collection methods for pebble count, macroinvertebrate samples, and habitat scoring data used in this study. ADEQ utilized a modified Wolman pebble count field method to gather particle size distribution data, and subsequently calculate percentage of fine sediment (total percent particles < 2 mm) and D50 in a stream reach (Wolman, 1954; Harrelson, et al., 1994). The pebble count is a direct measure of the median particle size (D50) for the reach or riffles sampled, and consists of measuring 100 particles at equal increments within the wetted width. From this data, the percent fine sediment (silt, sand, and clay) is calculated. Sediment data was collected from 1994 to 2006 in warm and cold perennial wadeable streams using the riffle pebble count method, in which only riffle habitats were sampled for particle size. Sediment measurements changed in 2007 when ADEQ adopted a reach-wide pebble count method for warm water streams, but retained the riffle pebble count also for cold water streams. Therefore, the riffle method not only applies to records from 1994 to 2006, but also to cold water streams from 2007 to 2016. The reach data only represents records post 2007, during which reach-wide pebble counts were collected for warm and cold water streams. This resulted in a 2007 to 2016 dataset containing riffle and reach-wide pebble counts for cold water streams and reach-wide only pebble counts for warm water perennial wadeable streams. Samples from effluent dependent waters, duplicate samples, streams affected by significant recent scouring, fires, or floods (year 2008), and large, nonwadeable river systems were excluded from this analysis. Sites are considered warm water below 5000 feet elevation, and cold water above 5000 feet elevation. Stream reaches were defined as reference, non-reference, or stressed site types a priori, based on whether general criteria were met, such as presence of upstream stressors or impacts (Jones, 2018). For fine sediment criteria, application of Arizona's bottom deposits numeric standard rule at A.A.C. R18-11-108(A)(1) is discussed in detail by Spindler in *Implementation Procedures for the Narrative Bottom Deposits Standard (2015)*.

Macroinvertebrate sampling was conducted during spring index period, from April to June. Samples were collected from representative riffle habitats only, taxonomic identifications were conducted by ADEQ contracted laboratories, and the data analyzed using ADEQ's Index of Biological Integrity.

Other variables used for analysis include habiat scoring measurements from ADEQ's Stream Ecosystem Monitoring (SEM) protocol, such as percent riffle, riffle and reach embeddedness, riparian canopy cover, percent algae and macrophyte cover, proper functioning condition of the riparian area (PFC), and Pfankuch channel stability index. Elevation and watershed area data were also included in the dataset.

Tables 1 and 2 list the sample sizes for each year and watershed included in the dataset. Variations are due to changes in sample plan, budget, and available field staff. A total of 555 Arizona stream sites were included in the final dataset (Figure 1). Of these sites, 340 were collected between 1994 and 2006, during which time ADEQ employed the riffle pebble count method, and 215 were collected between 2007 and 2016, using the reach pebble count method. The collection methods for the remaining habitat parameters varied slightly over time with the updating of SOPs, but were overall deemed to be comparable records. However, within the 555 site records, not all parameters were complete; holes in the data exist due to events such as incomplete sampling, data entry or transferal error, and changing sampling protocols. For this reason, statistical analyses had to be calculated either including the unavailable data (NAs), or only including records with all parameter values present.

SAMPLING AND ANALYSIS METHODS CONT.

Year	Number of Sites (n)
1994	68
1995	55
1996	17
1997	27
1998	24
1999	13
2000	32
2001	34
2002	17
2003	13
2004	19
2005	9
2006	12
2007	28
2009	29
2010	18
2011	14
2012	22
2013	40
2014	19
2015	21
2016	24

Watershed	Number of Sites (n)
Bill Williams	23
Grand Canyon	8
Little Colorado	108
Middle Gila	37
Salt River	83
San Pedro	74
Santa Cruz	13
Upper Gila	92
Verde River	117

Table 1 (left). Number of site samples collected by year. Table 2 (above). Number of site samples collected per watershed.

R statistical software through the RStudio platform was utilized in data analysis. Tests included Kruskal-Wallis and Wilcoxon tests for significance, Spearman coefficient correlations, and logistic regression.

STUDY AREA

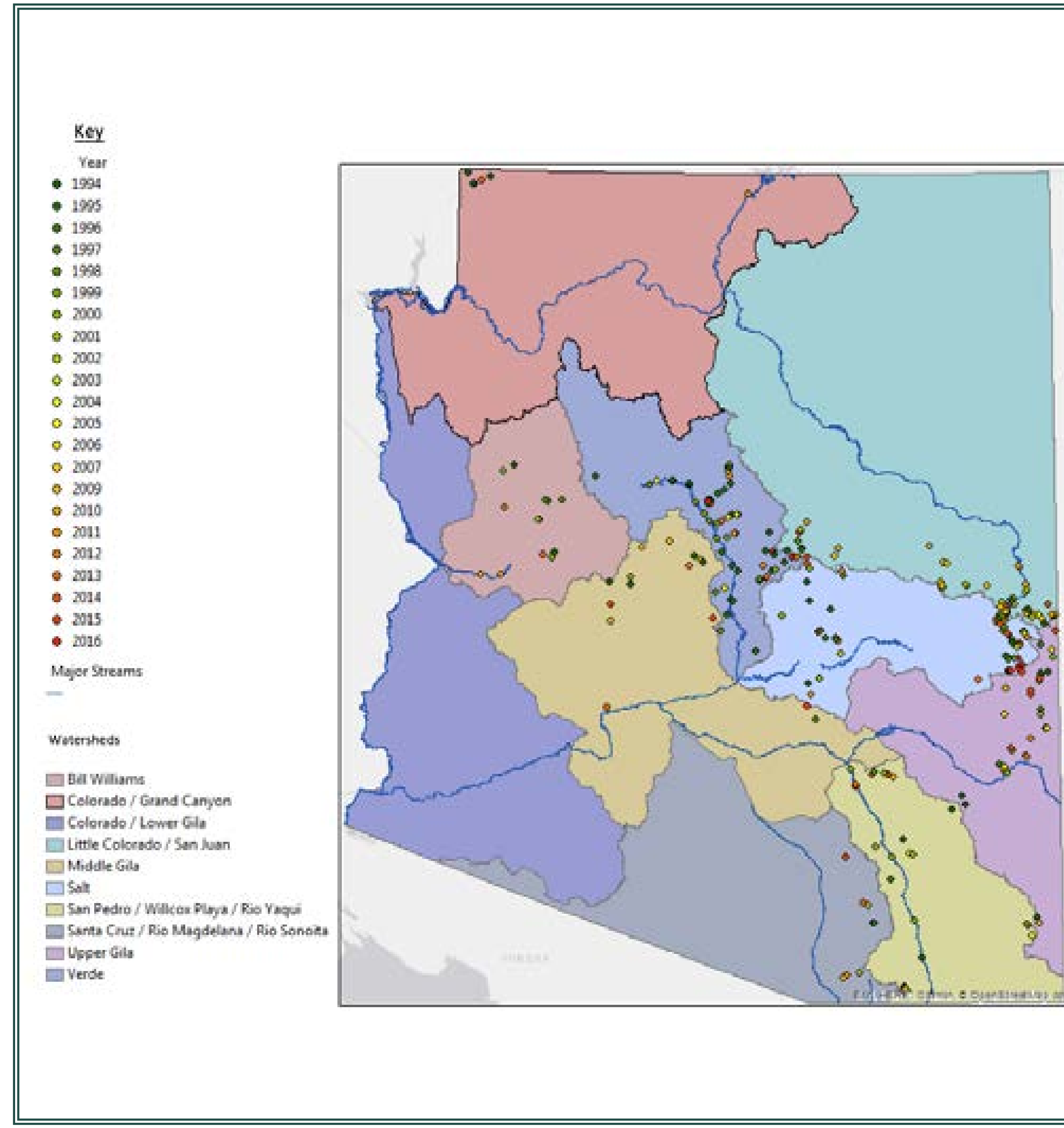

Figure 1. Map of Arizona with study sites plotted, organized by year of sampling, from 1994 to 2016, and watershed.

RESULTS

COMPARING PEBBLE COUNT METHODS

ADEQ riffle pebble count data (1994 to 2006) were compared with reach-wide pebble count data (2007 to 2016) in a boxplot analysis (Figures 2 and 3). The percent fines boxplot comparison shows the distribution of riffle data differs from that of the reach data. The median of riffle percent fines is lower than the median of reach percent fines; this difference is significant by the Wilcoxon test ($p < 0.05$, Figure 2). Similarly, the D50 particle size boxplot comparison of riffle and reach data shows different distributions. The median of riffle D50 data is significantly higher than the median of reach D50 data by the Wilcoxon test ($p < 0.05$, Figure 3).

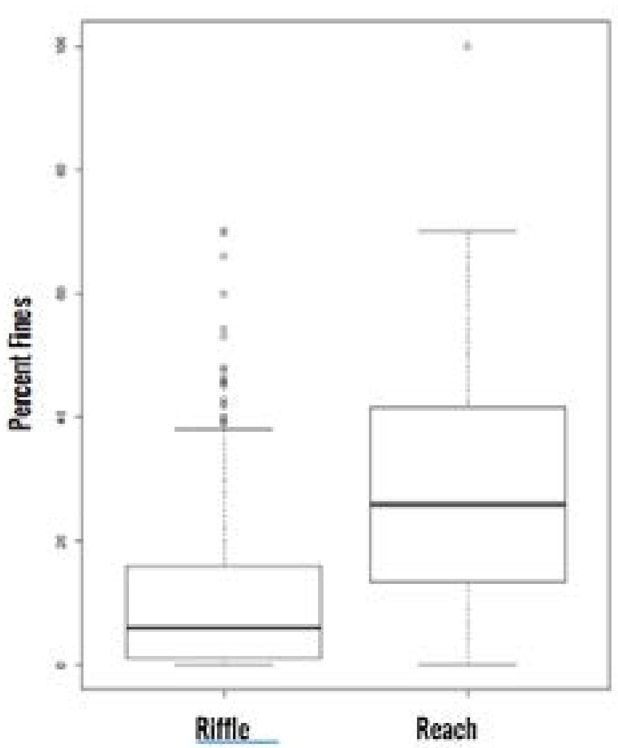

Figure 2 (left). Boxplots of percent fines data collected by riffle and reach methods (riffle n = 286; reach n = 199).

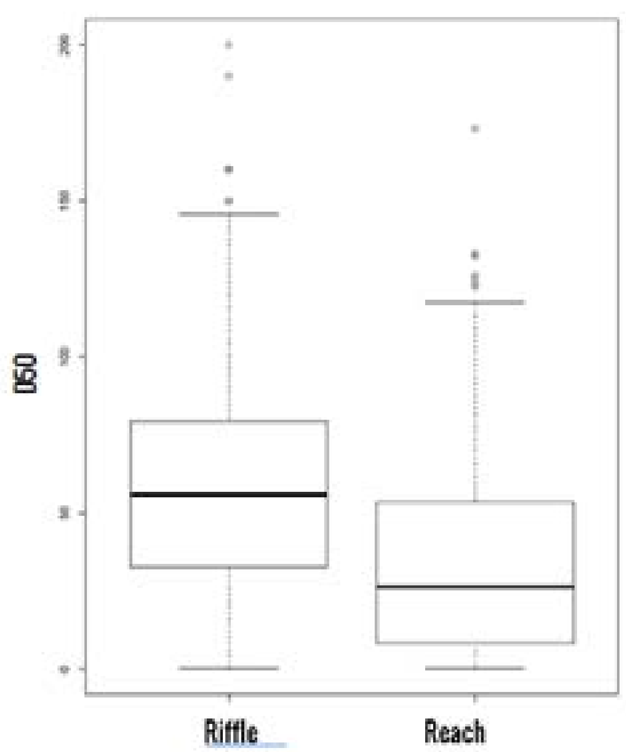

Figure 3 (right). Boxplots of D50 data collected by riffle and reach methods (riffle n = 286; reach n = 199).

The riffle method applies to records from 1994 to 2006 as well as to cold water streams from 2007 to 2016, as per ADEQ sampling SOPs, while the reach data only represents records for cold and warm water sites post 2007. This results in a larger riffle dataset. The riffle pebble count method yielded percent fines data that reflect low amounts of sediment in the streams sampled, and D50 particle size data that reflect larger average sediment size classes, compared to the reach pebble count method. The riffle method involves a predisposed selection of cobble habitats to sample over run and pool habitats, where fines aggregate as a result of sediment deposition patterns.

he 555 sites can be grouped by region based on elevation; warm water is less than 5000 feet and cold water is greater than 5000 feet, as previously stated. Each region can also be grouped by biocriteria rating, which relates the condition of the stream based on negative impacts from erosion, pollution, and possible anthropogenic sources to ecosystem health, based on relative abundances and diversity of benthic macroinvertebrate populations (Spindler, 2015a).

RESULTS CONT.

ADEQ riffle pebble count data (1994 to 2006) were compared with reach-wide pebble count data (2007 to 2016) in a boxplot analysis (Figures 2 and 3). The percent fines boxplot comparison shows the distribution of riffle data differs from that of the reach data. The median of riffle percent fines is lower than the median of reach percent fines; this difference is significant by the Wilcoxon test ($p < 0.05$, Figure 2). Similarly, the D50 particle size boxplot comparison of riffle and reach data shows different distributions. The median of riffle D50 data is significantly higher than the median of reach D50 data by the Wilcoxon test ($p < 0.05$, Figure 3).

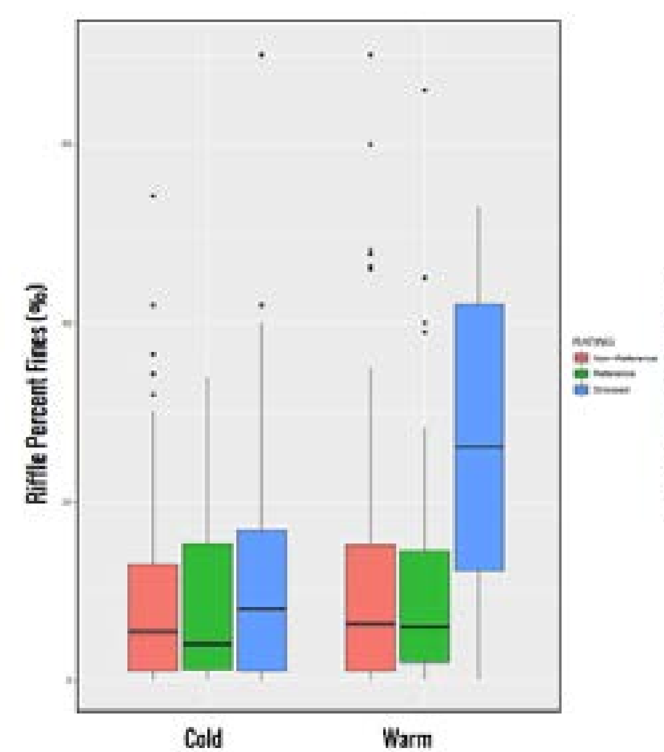

Figure 4 (left). Boxplots of riffle percent fines data grouped by site region (warm or cold) and site rating (reference, non-reference, or stressed). Sample sizes: cold-non-reference, n = 98; cold-reference, n = 34; cold-stressed, n = 39; warm-non-reference, n = 57; warm-reference, n = 48; warm-stressed, n = 10.

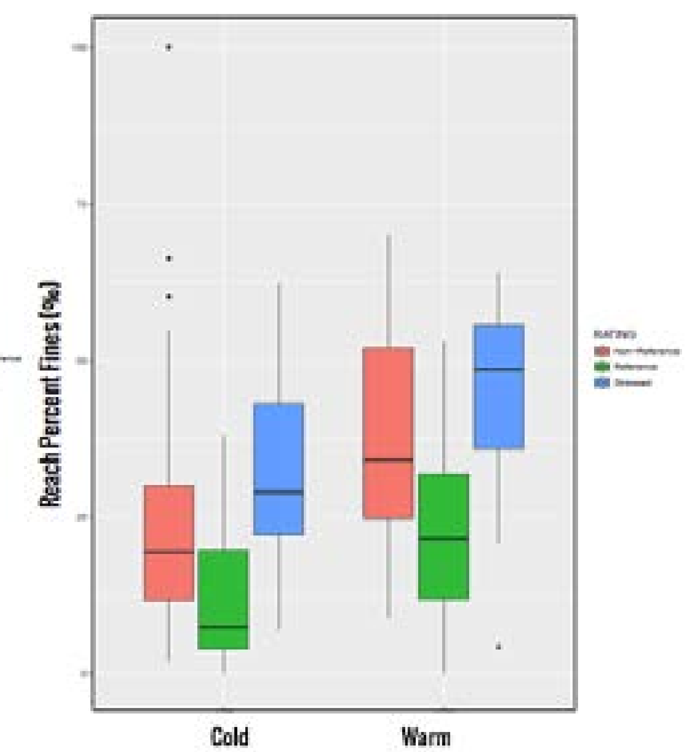

Figure 5 (right). Boxplots of reach percent fines data grouped by site region (warm or cold) and site rating (reference, non-reference, or stressed). Sample sizes: cold-non-reference, n = 64; cold-reference, n = 18; cold-stressed, n = 23; warm-non-reference, n = 48; warm-reference, n = 28; warm-stressed, n = 18.

A comparison of cold and warm site classes and rating categories produced some important results. Riffle percent fines data do not show a significant difference between reference and stressed distributions in cold water, meaning these site classifications portray similar habitat conditions of low fine sediment (Figure 4). Riffle percent fines in warm water data are significantly different between reference and stressed classifications (Figure 4). Reach percent fines data show a significant difference between reference and stressed data within both warm and cold regions (Figure 5). Reach fines data give more information about the difference between warm and cold region habitat conditions, given the distribution of warm water data has generally higher percent fines than the cold water data (Figure 5). It is important to note that the actual percent fines values available to generate plots for each group were limited, resulting in relatively small sample sizes, n (Figures 4 and 5).

RESULTS CONT.

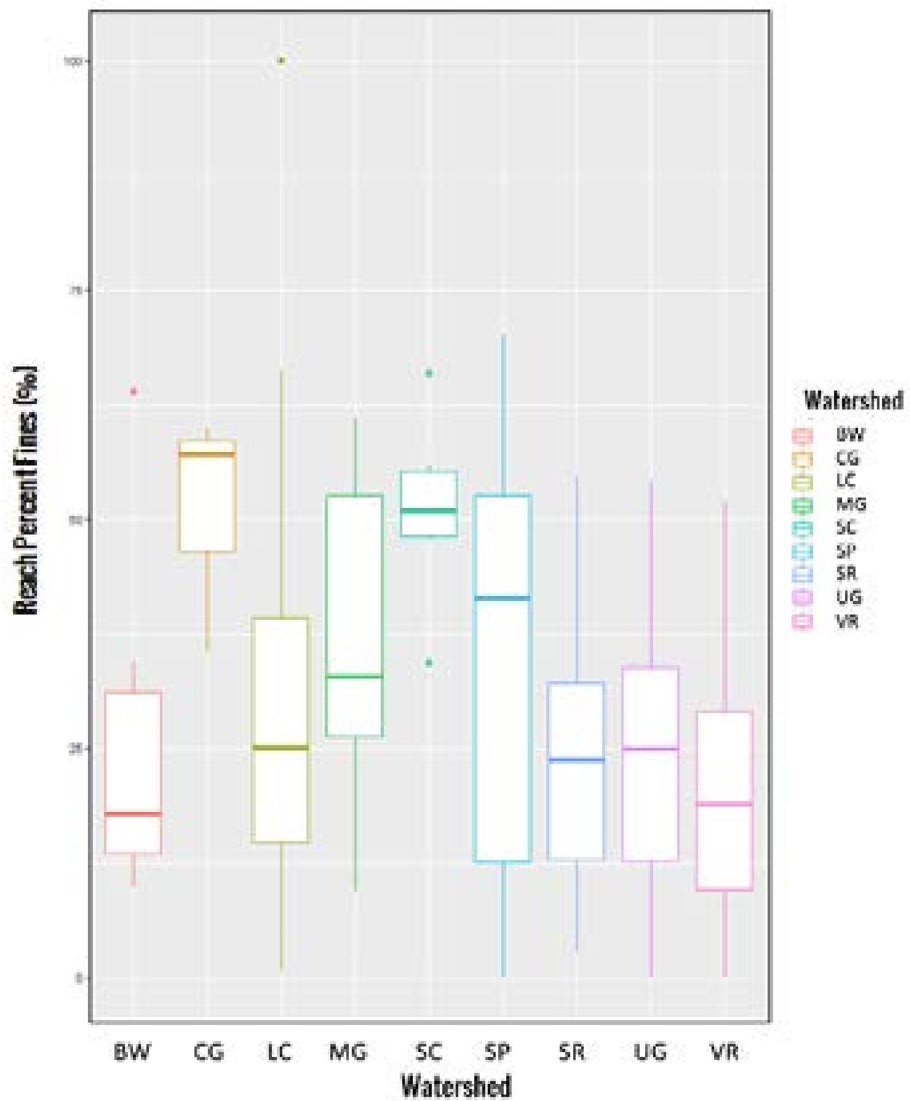

Figure 7. Boxplots of reach percent fines data grouped by watershed. Watersheds: Bill Williams (BW), Grand Canyon (CG), Little Colorado (LC), Middle Gila (MG), Santa Cruz (SC), San Pedro (SP), Salt River (SR), Upper Gila (UG), Verde River (VR).

Further exploratory analysis was performed to find associations of percent fines by watershed. The greatest reach-wide percent fines were found in the Santa Cruz and Grand Canyon watersheds, followed by the San Pedro and Middle Gila watersheds (Figure 7). The lowest quantities of reach-wide percent fines were found in the Verde and Bill Williams watersheds.

RESULTS CONT.

EMPIRICAL DATA COMPARISON WITH BOTTOM DEPOSITS STANDARD

The existing bottom deposits numeric thresholds, as developed from literature-based research methods, are 30% riffle fines for cold water streams and 50% reach fines for warm water streams (Table 3).

DESIGNATED USE (WADEABLE, PERENNIAL STREAM	BOTTOM DEPOSIT CRITERION (PERCENT FINE SEDIMENT <2 MM)
A&Wc	> 30%
A&Ww	> 50%

Table 3. Numeric bottom deposits standard thresholds (AAC R18-11-108.02).

In order to compare reach-wide pebble count percent fines data collected since 2007 to these standards, the site records were grouped by region and rating condition, and general statistics were calculated for each grouping (Table 4). The mean percent fines values for 5 of 6 site classes were less than the current bottom deposits standards. The median percent fines values were less than the current bottom deposits standards for all groups (cold-reference, cold-non-reference, cold-stressed, warm-reference, warm-non-reference, warm-stressed) (Table 4). The median values for the cold and warm water reference are 25% to 50% the value of the current bottom deposits standards (7.5% and 21.5%, respectively). The median values for the warm and cold stressed site groups are both less than the current bottom deposits standards (48.5% and 29.0%, respectively).

REGION	RATING	SAMPLES (N)	MEAN	STANDARD DEVIATION	MEDIAN	INTER-QUARTILE RANGE (IQR)
Cold	Reference	18	13.1	12.1	7.5	15.8
Cold	Non-Reference	64	23.3	17.7	19.5	18.3
Cold	Stressed	23	32.0	15.7	**29.0**	20.9
Warm	Reference	28	23.5	15.6	21.5	20.0
Warm	Non-Reference	48	37.5	18.2	34.2	27.3
Warm	Stressed	18	44.3	15.5	**48.5**	19.8

Table 4. Data summary of reach percent fines data for study sites grouped by region and rating.

The values presented in Table 4 refer to reach-wide pebble count method calculations of percent fines for both warm and cold water sites, because it is the more representative method of stream conditions. Other data of interest are the riffle percent fines calculated for the cold water sites. Median fines values for reference (4.0%), non-reference (5.5%), and stressed (8.0%) sites also are much less than the 30% riffle fines standard for cold water sites. Given that riffle counts tend to underestimate sediment conditions, as mentioned above, the riffle fines values show a greater disparity with the current bottom deposits standard than reach fines.

RESULTS CONT.

BIOLOGICAL CORRELATIONS

The Arizona macroinvertebrate Index of Biological Integrity (IBI) was developed by ADEQ to provide a framework for assessing the biological condition of warm and cold water perennial, wadeable streams. These bioassessment indexes are based on various attributes, such as taxa richness, based on a list of taxa and their abundances (Spindler, 2006). The cold water IBI score and A&W use is meeting standards when the value is ≥ 52 and the warm water IBI score and A&W use is met when the value is ≥ 50 (Table 5; Spindler, 2015b).

MACROINVERTEBRATE BIOASSESSMENT SAMPLE RESULTS	INDEX OF BIOLOGICAL INTEGRITY SCORE		
	Cold water	**Warm water**	**Status**
Greater than the 25th percentile of reference condition	≥ 52	≥ 50	**Meeting**
Greater than the 10th and less than the 25th percentile of reference condition	46 – 51	40 – 49	**Inconclusive**
Less than the 10th percentile of reference condition	≤ 45	≤ 39	**Violating**

Table 5. Arizona Index of Biological Integrity thresholds for wadeable, perennial streams with aquatic and wildlife cold (A&Wc) or warm (A&Ww) designated use (Spindler, 2015b).

IBI scores were used in this dataset to measure the biological response to percent fines and other habitat parameters collected as per ADEQ protocols referenced above. These habitat parameters included riparian vegetation canopy density, percent fines by reach method, D50 by reach method, percent cover of filamentous algae, percent cover of macrophytes, reach percent embeddedness, riffle percent embeddedness, percent ideal Proper Function Condition (PFC) score, percent ideal Pfankuch score, percent ideal habitat score, percent riffle, elevation, and watershed area. Cold water data includes D50 and percent fines by riffle method, as well (Figure 8). Warm water data only include reach values because riffle pebble count was only performed at cold water sites post 2007.

Correlation tests were conducted to determine which other habitat parameters with which percent fines might be related, and to determine if percent fines or other habitat parameters correlate best with IBI score. Spearman correlations confirm that habitat factors such as D50, percent fines, and embeddedness are directly correlated to each other in warm and cold regions (Figures 8 and 9). This is an expected correlation, as they are all measurements of sediment in the stream bottom substrate. The riparian metric, PFC, channel stability metric, Pfankuch, and substrate habitat scores are all correlated in both warm and cold water datasets. In the cold water region dataset (Figure 8), IBI score is positively correlated with canopy density (0.34) and elevation (0.53), and negatively correlated with percent filamentous algae cover (-0.34), reach embeddedness (-0.27), and watershed area (-0.47). In the warm water dataset (Figure 9), IBI score is positively correlated with habitat score (0.28), and negatively correlated with reach embeddedness (-0.21).

RESULTS CONT.

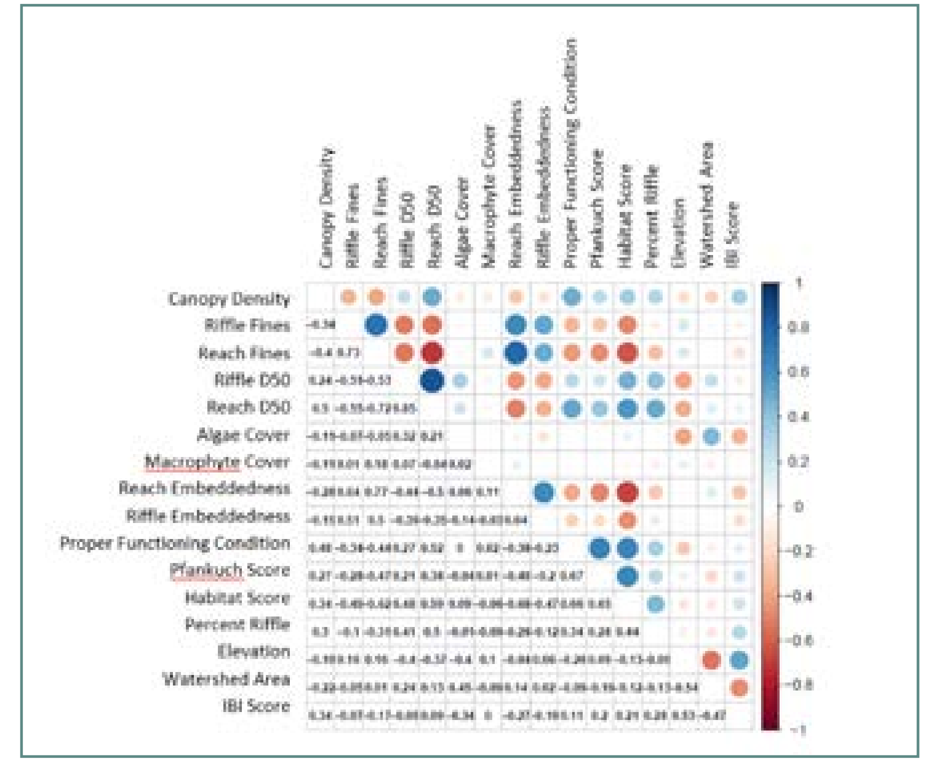

Figure 8. Spearman correlation plot of cold water data comparing each observed variable.

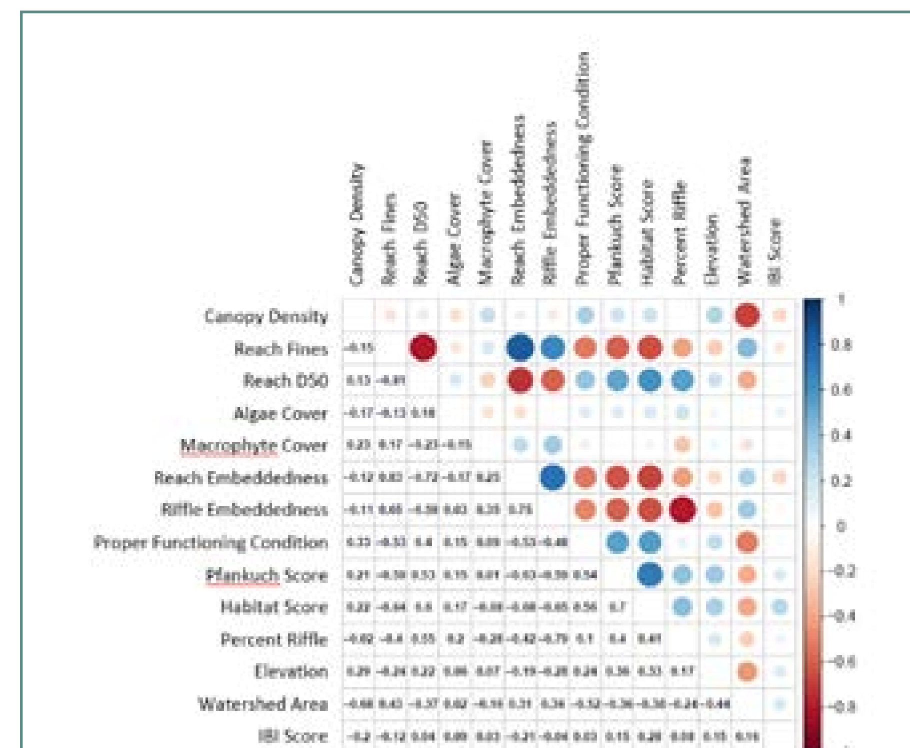

Figure 9. Spearman correlation plot of warm water data comparing each observed variable.

RESULTS CONT.

The expected correlation of IBI with percent fines or D50 values was not observed to a significant degree. Both warm and cold water data showed weak negative correlation coefficients of IBI score to reach fines (-0.15 and -0.40, respectively), and weak positive correlation coefficients of IBI score to reach D50 (0.13 and 0.50, respectively). The biotic index (IBI) is not significantly correlated with any other habitat variable included in this study, according to the Spearman correlation test (Figures 8 and 9). A follow up analysis was performed using recent data only (post 2007) to determine if removing the outdated field protocol results would reduce possible noise in the tests, but this had only a minimal effect on IBI score correlations.

Since the IBI score is a blunt tool that is not specifically tuned to sediment conditions, a logistic regression method was used to research the relationship between excessive sediment and biological response by macroinvertebrate indicators. In a previous ADEQ study, sediment weighted averages were calculated based on relative abundances of macroinvertebrate taxa and percent fines values from 143 cold water and 189 warm water sites collected from 1997 to 2007 (Spindler et al., 2008). Several taxa were identified with weighted averages that indicated high or low tolerance to fine sediment. The taxa in Table 4 each represent a genus of aquatic macroinvertebrate that is either considered sensitive or tolerant to sediment, as determined by the 2008 ADEQ study. Two sensitive and two tolerant taxa were chosen for both warm water and cold water for this logistic regression test, with abundance data converted to presence/absence binary data (1/0). Logistic regression was performed using R statistical software to determine whether sediment level, quantified as reach percent fines, affects taxa presence. The binary dependent variable was presence/absence of the selected taxa, using percent fines and D50 as explanatory variables, to create models for probability of presence of each genus given increasing percent fines. Model fit was assessed by completing a multi-collinearity test of the variables, using the R "vif" function. All models for the taxa were assessed as having low instance of confounding collinear data.

	Taxa	Effect Coefficient	P-value
Sensitive	Brachycentrus[1]	-0.0409	0.236
	Epeorus[1]	-0.0417	0.0162
	Serratella[2]	-0.0376	0.0797
	Polycentropus[2]	-0.0743	0.00362
Tolerant	Cheumatopsyche[3]	0.00897	0.416
	Tricorythodes[1]	0.0102	0.381
	Tanytarsus[2]	0.00295	0.786

Table 6. Results from logistic regression analysis using percent fines and D50 data as continuous explanatory variables to test against binary presence/absence data for the eight taxa.
[1] cold water genera; [2] warm water genera; [3] warm and cold water genus

RESULTS CONT.

The results of the logistic regression revealed that sediment-sensitive taxa for warm and cold water had more significant correlation with percent fines than tolerant taxa, given p-values for Brachycentrus, Epeorus, Serratella, and Polycentropus, compared to those for Cheumatopsyche, Tricorythodes, and Tanytarsus (Table 6). Further, the sensitive taxa expressed negative effect coefficients (odds ratios) with sediment as explanatory variables, while tolerant taxa had lower magnitude, positive effect coefficients. It can be interpreted that the probability of presence for sediment-sensitive genera is significantly negatively affected by percent fines, but sediment-tolerant genera presence is not reliant on sediment level.

There are several limitations associated with this data. The model was not able to suggest a discrete level of percent fines at which the taxa are most responsive. Because logistic regression requires data with no missing values, records with gaps in values reported had to be excluded. This left each set of taxa data with low "n" sample sizes. In conjunction with this, the taxa dataset includes rare "events," meaning instances of taxa presence are few compared to instances of absence. In addition, when the complete dataset is assessed separately as warm versus cold region data to correlate with warm- or cold-associated taxa, the sample size is further decreased. As a result, the final sample sizes, despite initial n = 555, representing presence of each taxa in warm water and cold water sites were as low as n = 3 (Polycentropus, warm) and n = 4 (Brachycentrus, cold). To resolve this issue, logistic regression models were built using warm and cold datasets combined to increase sample size and improve fit despite the rarity of events. This data should therefore be considered preliminary, but also considered as an implication of a significant relationship between sediment and biotic response. Further analysis of this data should be conducted to include exact logistic regression, a model which accounts for rare events (1), or presence, among frequent nonevents (0), or absence. Another viable option is to include analysis of other sensitive taxa suggested from literature (Relyea et al, 2011).

DISCUSSION

The primary objectives of this study were to: 1) evaluate the existing bottom deposit standards thresholds, using ADEQ datasets to determine whether the existing criteria need to be updated, 2) conduct exploratory analyses on ADEQ datasets to determine the relationships between stream sediment parameters and macroinvertebrate indicators, and 3) determine whether current sampling methods adequately represent stream sediment conditions.

The dataset was reduced from tens of thousands of records, to only include 555 samples, 215 of which were collected in recent years (2007-2016, 10 years). The older records (1994-2006, 13 years) contain data based on outdated habitat assessment methods, especially the riffle count procedure. This underscores the restrictions of this dataset due to method changes and need for larger sample sizes. The riffle pebble count method significantly underestimates percent fines and overestimates D50 distributions in the stream compared to the reach pebble count method. This is due to the predisposed selection of cobble habitats over run and pool habitats, where fines aggregate as a result of sediment deposition patterns. Therefore, data collected between 1994 and 2007, as well as cold water riffle data since 2007, has underrepresented stream conditions for bottom deposits. Overall, cold water reach data have lower percent fines values than warm water reach data. At cold water sites higher in elevation, and therefore higher in the watershed, sediment is eroded and carried downstream. At sites lower in the watershed, deposition of sediment is a natural process, resulting in higher fines in the warm water reaches. Reach fines data is more representative of this process than riffle fines, given the distributions of data.

The reach pebble count method should be used over the riffle method for collecting percent fines and D50 data. The reach percent fines data imply that the reach pebble count method is the better protocol for obtaining a representative sample of the stream conditions, compared to the riffle method. The riffle method biases results to include only cobble habitat, thus creating left-skew in the data to favor lower percent fines. This in turn provides a misguided interpretation of the stream sediment load, and will underrepresent the possible impairments present in the state. Similarly, riffle sampling of benthic macroinvertebrates, as described in the ADEQ sampling methods, limits the biological sample by neglecting to sample the possible multihabitat zones, such as pools and runs, which may be favorable to different species (Jones, 2012). Methods exist for determining the appropriate macroinvertebrate sampling method based on stream ecosystem type, such as limestone streams, multihabitat pool/glide streams, and riffle/run freestone streams (Botts, 2009). A similar sampling scheme would provide a robust system to emulate for truly representative macroinvertebrate sampling of Arizona streams.

Arizona has diverse ecoregions, as a function of the wide ranges of elevation and biomes in the state. ADEQ has attempted to parse out the variation in the data by using cold/warm region classifications based on site elevations. To reduce the impact of outliers on the dataset, study sites can be classified further by watershed or hydrogeomorphic stream type, to compare the natural size range of streambed sediments with the stream erosive competence (Rosgen, 1994; Kaufmann et al., 2007). Using ecoregions and erodibility factors for site classification structure, and relative bed stability for more robust sediment calculation, the response of macroinvertebrates can be more accurately assessed (Jessup et al., 2014). Assuming predisposed variability in percent fines data, such as demonstrated in Figure 7 which categorizes by watershed, may improve the distribution of sediment data and reveal more effective analysis for determining stream condition. These stream systems tend to carry naturally higher sediment loads. This is important when assessing whether a stream, which may be exceeding bottom deposit standards, is impaired by sediment or is representative of natural sediment conditions. The stream habitat may be healthy and meeting all standards aside from naturally high fines. Another objective for this data may be to determine the benefit of developing sediment standards related to ecoregion, as well as the current elevation-based biocriteria regions. This option is explored by Relyea et al. in a New Mexico study that established different reference conditions based on mountains, foothills, and xeric ecoregions (Relyea et al., 2011). The study was able to define a significant biological response to the recommended benchmarks; this could be an important next step for improving ADEQ biocriteria rules for future review periods.

At current standards, the medians of all site types (reference, non-reference, and stressed, for cold and warm water) meet bottom deposits criteria, as listed in A.A.C. R18-11-108(A)(1). Of the 298 warm water sites included in this study, 25 exceeded the 50% fines threshold for reach data; of the 257 cold water sites, 15 exceeded the 30% fines threshold for riffle data. The standards are not stringent enough to distinguish all sites that should be considered stressed or impaired based

DISCUSSION CONT.

on ADEQ habitat assessments and pebble count data. This study recommends that standards be based on reference data thresholds, to determine the point above which a stream ecosystem is considered healthy and functional under certain sediment conditions. These values reflect the 2012 Report "A Statewide Assessment of Arizona Streams," in which ADEQ developed Arizona reference site-based criteria for cold and warm water streams for the purpose of assessing streams within the report only. Reference values constituting "least disturbed" conditions were <16% for cold water and <20% for warm water. These values relate closely to the results in Table 4. The median percent fines value for cold water reference site data is even lower than the 2012 report value, meaning it even further underscores the current bottom deposit standard. The median for warm water data is very comparable to the value determined by the 2012 report, reiterating the need for an updated threshold based on reference condition data. Because this study incorporates data up to 2016, while the 2012 assessment report used data up to 2009, we can assume that the more recent dataset follows similar trends as the dataset analyzed previously. This speaks to the possibility of using these consistent trends as reference points to generate new bottom deposit numeric criteria.

Macroinvertebrates have a sensitive response to many variables aside from fine sediment, vegetation cover, and habitat scoring measurements. Literature provides evidence for the relationship between excessive bottom sediment in streams and the associated negative impact on biological variables. Benthic macroinvertebrates are a well-known bioindicator of sediment stress, exhibiting community shift at high levels of fines (Relyea, et al., 2011). However, a stream may be stressed by one or several of many other habitat or chemical factors, such as metal exceedances, and these stressors can affect the macroinvertebrate community, apart from excessive fine sediment. The response of macroinvertebrates to their environment is nuanced; as a generalized calculation of population and diversity, IBI score may be masking these minute processes and responses.

IBI scores and percent fine sediment did not have strong direct relationships; however, not all parameters outside of physical habitat variables that affect macroinvertebrate populations are included in the correlation analyses (Figures 8 and 9). Also, IBI score is composed of several metrics and is somewhat of a blunt tool, not tailored to respond specifically to sediment stressors, but rather to all stressors. The index of biotic integrity (IBI) score will instead need to be correlated to relative bed stability (RBS), geomorphology, and chemistry data, as well as habitat parameters (Relyea et al., 2000). Macroinvertebrates are sensitive to many factors aside from sediment, and respond by shifting community structure and relative abundances of genera and species toward those more tolerant to the stressors (Brusven and Prather, 1974). The next step is to follow up with calculating weighted averages of bioindicator species, such as those sensitive to sediment (Table 4) using recent data, post 2008, to analyze how the community responds to habitat parameters, on a level more specific than IBI score. Fine sediment thresholds for warm and cold regions will be determined by indicator species, using abundance percentiles of sediment-sensitive and sediment-tolerant species. Thresholds for bottom deposit criteria development may also be developed by continuing the logistic regression analysis as presented in this study, in order to determine the level of percent fines at which the probability of presence of indicator macroinvertebrates most severely declines.

Fish spawning habits and juvenile fish stages are another sensitive indicator of excessive fine sediments in streams. It has been demonstrated that fish show effects of fine sediment impeding oxygen exchange in gravel spawning beds and suffocating young (Waters, 1995; Chapman and McLeod, 1987; Bjornn et. al., 1977). The development of bottom deposits criteria may also be based on these biological effects, in conjunction with sensitive macroinvertebrate taxa response. The inclusion of fish data will provide more corroborating data on which the new thresholds may be developed, and improve the protection of aquatic life.

Based on the historic dataset collected by ADEQ from 1994 to 2016, this study was not able to definitively calculate updated bottom deposits criteria to replace the current literature-based thresholds. However, the presented analyses reveal the extent to which the thresholds are not representative of Arizona streams, and may be causing Type II error in impairment decisions for streams stressed by elevated sediment conditions. More reach-wide habitat data will need to be collected, with further analysis using exact logistic regression and weighted averages of sensitive taxa, in order to revise the standard for tentative submission during the next review period. Developing a new bottom deposits standard that will better protect aquatic life in Arizona will serve to continuously improve environmental regulation.

PAGE INTENTIONALLY LEFT BLANK

REFERENCES

ADEQ. 2003. Arizona Administrative Code, Title 18, Chapter 11. Water Quality Standards for Surface Waters. Arizona Department of Environmental Quality, Phoenix, AZ.

Bjornn, T.C., M.A. Brusven, M.P. Molnau, and J.H. Milligan. 1977. Transport of granitic sediment in streams and its effects on insects and fish. University of Idaho, College of Forestry, Wildlife and Range Sciences, Bulletin Number 17. Moscow, Idaho.

Botts, W. 2013. Instream Comprehensive Evaluation Surveys. Bureau of Clean Water, Pennsylvania Department of Environmental Protection. Harrisburg, PA.

Brusven, M.A. and K. Prather. 1974. Influence of stream sediments on distribution of macrobenthos. J. Entomological Society B.C. 71:25-32.

Chapman, D.W. and K.P. McLeod. 1987. Development of criteria for fine sediment in the northern Rockies ecoregion. U.S. Environmental Protection Agency, USEPA Report 910/0-87-162, Washington, D.C.

Jones, J.D. ed. 2012. Standard Operating Procedures for Surface Water Quality Sampling. Arizona Department of Environmental Quality, Surface Water Section Technical Bulletin, TB06-02. Phoenix, AZ.

Jones, J.D. and Woods, J. ed. 2012. A Statewide Assessment of Arizona's Streams: July 2007 – June 2010. Arizona Department of Environmental Quality, Surface Water Section, OFR13-01. Phoenix, AZ.

Kaufmann, P.R., J.M. Faustini, D.P. Larsen, and M.A. Shirazi. 2007. A roughness-corrected index of relative bed stability for regional stream surveys. USEPA. Elsevier: Geomorphology. 10.1016.

Relyea, C.D., G.W. Minshall, and R.J. Danehy. 2000. Stream Insects as bioindicators of fine sediment. Idaho State University, Department of Biological Sciences, Stream Ecology Center. Pocatello, Idaho.

Rosgen, David L. 1994. A Stream Classification System. Catena. Volume 22. Pp 169-199. Elsevier Science, Amsterdam.

RStudio. 2016. RStudio: Integrated Development for R. RStudio, Inc, Boston MA. URL – http://www.rstudio.com/.

Spindler, P.H. 2004. Stream Channel Morphology and Benthic Macroinvertebrate Community Associations in the San Pedro River and Verde River Basins of Arizona. Arizona Department of Environmental Quality. Phoenix, AZ.

[a]Spindler, P.H. 2015. Implementation Procedures for the Narrative Biocriteria Standard. Arizona Department of Environmental Quality. Phoenix, AZ.

b Spindler, P.H. 2015. Implementation Procedures for the Narrative Bottom Deposits Standard. Arizona Department of Environmental Quality. Phoenix, AZ.

Spindler, P.H., A. Condon, J. Jones, J. Woods, M. Smart, K. Ehrenberg, L. Johnson, L. Lawson, K. Elrod, K. Hanson, and J. Bern. 2008. Fine sediment tolerance in benthic macroinvertebrates of Arizona. Society for Freshwater Science Poster Presentation. Arizona Department of Environmental Quality. Phoenix, AZ.

USEPA. 2002. National Water Quality Inventory 2000 Report. U.S. Environmental Protection Agency Report #EPA-841-R-02-001. Washington, D.C.

Waters, T.F. 1995. Sediment in streams: sources, biological effects and control. American Fisheries Society Monograph 7. Bethesda, MD.

Wolman, M.G. 1954. A method of sampling coarse river-bed material. American Geophysical Union Transactions, (35):951-956.

PAGE INTENTIONALLY LEFT BLANK

APPENDICES

APPENDIX A

HABITAT PARAMETERS

Sample of data table of compiled habitat parameters for each of 555 stream sample sites in Arizona. Please contact author for full data table.

STATIO	REGION	WS	DATE	LAT	LONG	CNPY_	FINES_	FINES_	D50_R	D50_R	PFA	AVP	EMBED_	EMBED	PFC_PCT	PFANKUC	HABIT	RIFFLE	ELEV_FT	WS_AREA	IBI	RATING	Brachy	Cheum	Epeor	Polycd	Serrat	Tanyta	Tricory
MGAF R087.86	warm	MG	4/6/2009	34.314	-112.061	26		54.7		1.4	12.5	12.5	57.255		80	67.7419	40	1.6	3445	587.47	20.32	Stressed						3	7
MGH SR00 2.67	warm	MG	5/16/2013	33.347	-112.726	9.5		35		28.5	0.5	0.5	68.72		13.333	50.5376	45	53.1	831.9	1471	24.31	Stressed							
CGPA R000.40	warm	CG	4/24/2013	36.864	-111.597			60		123	0.5	12.5	79.12		12.5	29.0323	47.5	28.6	1000	1417.55	25.48	Stressed							
SPSP R015.43	warm	SP	4/3/2012	32.813	-110.698	35		67.6		0.2	37.5	12.5	77.89		56.25	48.3871	50	30.4	2190	1168	25.65	Non-Reference						3	1
VRVE R185.31	warm	VR	5/9/1995	34.867	-112.402												61		4200	2487.55	28.26	Stressed							1
VRFV R034.80	warm	VR	4/21/2003	34.299	-111.358	11.6	1		65		0.5	0.5		26			67.5	16	4500	140.9	28.54	Non-Reference						13	
CGVG R044.58	warm	CG	4/24/2013	36.921	-113.861	0		57.1		1.3	37.5	12.5	71.9		46.667	72.043	55	26	1970	4194	28.79	Stressed							
VREV R034.80	warm	VR	5/3/2001	34.299	-111.358	46.5	1		61			12.5		55			80	26	4500	140.9	30.38	Non-Reference						26	
SRCH E013.55	warm	SR	5/8/1996	33.828	-110.857												52		3190	199	31.6	Reference		19			7		
MGAF R110.91	warm	MG	4/14/2015	34.49	-112.231	58.8		31		9.5	1	29	61.21	38.5	93.75	77.4194	80	19.2	4410	174.64	32.74	Stressed							74
VRCM P019.06	warm	VR	4/28/1998	33.879	-111.803	46	15		14		12.5	12.5					80	58	2960	11.92	33.44	Non-Reference							
SCCE 006.60	warm	SC	4/17/2001	32	-110.602	55.5	10.5		13		37.5	0.5		71.2			35	3.5	3491	341	34.18	Non-Reference							18
SPSP R015.43	warm	SP	4/3/2013	32.813	-110.698	40		70		0.5	12.5	62.5	80.2		52.941	55.914	40	18.4	2190	1168	34.47	Non-Reference							25
SCSO N014 53	warm	SC	4/13/2001	31.509	-110.802	60	1		52		0.5	12.5		36			72.5	52.4	3880	198.14	34.57	Stressed							
MGA NT01 3.35	warm	MG	4/29/1998	34.196	-112.715		15.7		25		87.5	12.5					70	56	3850	5.18	34.99	Non-Reference		1					
VRPIP 008.10	warm	VR	5/2/2013	34.319	-111.458	78		4.3		108	62.5	0.5	35.3	13	100	80.6452	95	34.6	4375	32.05	35.42	Stressed			1				

Arizona Department of Environmental Quality

www.azdeq.gov
July 2018
Publication Number: OFR-18-03