Sediment in New Mexico Streams: Existing Conditions and Potential Benchmarks

New Mexico Environment Department
United States Protection Agency
Tetra Tech, Inc.

Final Report
August, 2010

Sediment in New Mexico Streams:
Existing Conditions and Potential Benchmarks

Prepared for:

New Mexico Environment Department
PO Box 26110 - 1190 St. Francis Drive N4050
Santa Fe, New Mexico 87505

and

U.S. Environmental Protection Agency
Region 6
1445 Ross Avenue, Suite 1200
Dallas, TX 75202-2733

Prepared by:

Tetra Tech, Inc.
Center for Ecological Sciences
73 Main Street, Suite 38
Montpelier, VT 05682

Final Report
August 2010

Executive Summary

Purpose

The State of New Mexico Environment Department (NMED) associates imbalanced suspended and bedded sediment supply with effects to aquatic life uses through narrative or comparative standards (New Mexico Administrative Code 20.6.4.13). The degrees to which certain sediment quantities are unnatural and detrimental to associated aquatic life are as yet loosely defined. The U.S EPA Region 6 and NMED are interested in developing benchmarks by site class to better implement the existing narrative criteria.

The purpose of these analyses is to identify sediment characteristics that are expected under the range of environmental settings in New Mexico, especially in undisturbed reference streams. Through this characterization, it will be possible to identify situations where the expectations are not met, using sediment indicators that show responsiveness to disturbance. Associating biological measures with sediment indicators will further indicate situations where the disturbance causes sediment imbalance and biologically-relevant habitat degradation. The results of these analyses will allow recommendations regarding the application of quantitative sediment benchmarks on New Mexico perennial streams.

Methods

The approach to setting sediment benchmarks in New Mexico followed seven basic steps:

1. Review background information
2. Assemble datasets
3. Establish reference sites
4. Classify sites
5. Characterize sediments
6. Describe stressor-response relationships
7. Recommend benchmarks

The analytical approach was to corroborate sediment expectations and sediment-biological relationships through multiple techniques. Three analytical techniques were used to identify potential benchmarks for sediments. These include analyses of reference distributions, quantile regression, and change-point analysis. Reference distributions describe expectations in least disturbed sites, which are classified by natural site types, as needed. Quantile regression and change-point analysis compare sediment conditions with biological conditions, using the biological conditions to indicate the degree to which aquatic life uses are supported. This report presents the results of each analysis and recommends possible benchmarks for NMED to consider. Selection of final sediment benchmarks using the weight-of-evidence from multiple analytical approaches will provide a solid basis for protective management strategies.

Results

Analysis of multiple sediment indicators, their responsiveness to site disturbance, and their effects on benthic macroinvertebrates resulted in identification of potential benchmarks for the bedded sediment indicators % sand & fines and Relative Bed Stability (LRBS_NOR) in three site classes, Mountains, Foothills, and Xeric areas. The site classes distinguish sediment expectations across the State and were identified through a principal components analysis (PCA) of environmental conditions and the sediment indicators. Percent sand & fines are easily measured and related strongly with biological metrics. LRBS_NOR is a formulation that considers site-specific hydraulic potential for moving bed sediments, so that the observed fine sediments are only considered imbalanced when the streambed is more easily mobilized and transported than expected due to unstable stream bed conditions. The two indicators can be applied in a two-tiered assessment that first considers the simpler indicator of biological impairment, and then refines the assessment with the second indicator of geomorphic impairment, as needed. Recommended benchmark values are as follows:

Site Class	% sand & fines	LRBS_NOR units
Mountains	< 20	> -1.1
Foothills	< 37	> -1.3
Xeric	< 74	> -2.5

In the Xeric sites, given the high % sand and fines benchmark, an alternative benchmark can be considered for % fines, where > 29% would indicate stress. This benchmark is also based on biological responses.

Suspended sediments were most commonly measured during low-flow conditions, when they are least stressful to benthic macroinvertebrates. Low-flow suspended sediment measures were not strongly related to biological metrics, but they were related to high-flow sediments, which were in turn related to biological metrics and bedded sediment measures. However, data were not sufficient for identifying a biologically-based low-flow or high-flow benchmark. Potential benchmarks were based instead on the distributions of values in sites that were fully supporting their aquatic life uses. Given a lack of adequate data and concerns expressed by NMED staff with using impairment support decisions as the basis for the distribution analyses, specific benchmarks, while identified, cannot be recommended for assessment of suspended sediment at this time.

Conclusions

Benchmarks for bedded sediments were established using multiple methods that yielded congruent results. Therefore, the recommended benchmarks and supporting analyses provides a strong basis for NMED to make final selections of benchmarks and to establish procedures for their application.

Suspended sediment benchmarks were derived from the distributions of values in sites that were fully supporting their aquatic life uses. The low-flow measures were not strongly related to benthic macroinvertebrate metrics. However, the relationships observed between low-flow and high-flow suspended sediments, high-flow sediments and biological responses, and suspended and bedded sediments all support the assumptions that low-flow suspended sediments are a valid indicator of habitat suitability for aquatic fauna. Patterns observed among site classes were credible in light of the characteristics of sites in the bedded sediment data set. Additional data and analyses are needed to be able to recommend specific benchmarks for suspended sediment.

Acknowledgements

This report was prepared with considerable input and feedback from a number of individuals. The New Mexico sediment workgroup included the following:

James Hogan, NMED
Lynette Guevara, NMED
Seva Joseph, NMED
Doug Eib, NMED
Forrest John, USEPA Region 6
Phil Kaufmann, USEPA Western Ecology Division
Angel Kosfiszer, USEPA Region 6
Benjamin Jessup, Tetra Tech, Inc.

Additional technical support was contributed by Shann Stringer of NMED (former) and Tetra Tech (current), Erik Leppo, Nick Jokay, and Brenda Decker of Tetra Tech, and Robert Kirkland, USEPA Region 6. John Olsen of the Utah State University provided information on the relative hardness of geologic units. Cover photos were provided by NMED.

Recommended citation:
Jessup, B.K., D. Eib, L. Guevara, J. Hogan, F. John, S. Joseph, P. Kaufmann, and A. Kosfiszer. 2010. Sediment in New Mexico Streams: Existing Conditions and Potential Benchmarks. Prepared for the U.S. Enviromental Protection Agency, Region 6, Dallas, TX and the New Mexico Environment Department, Santa Fe, NM. Prepared by Tetra Tech, Inc., Montpelier, VT.

Table of Contents

Appendices

List of Tables

List of Figures

Abbreviations

%Fines	Percent Fines
%SaFn	Percent Sand and Fines
ADEQ	Arizona Department of Environmental Quality
ALU	Aquatic Life Use
ANOVA	Analysis of Variance
CDPHE	Colorado Department of Environment and Public Health
CI90	90% confidence interval
cm	centimeter
CW	cold water
DE	Discrimination efficiency
DBH	Diameter Breast Height
EDAS	Ecological Data Application System
EMAP	Environmental Monitoring and Assessment Program
EPA	Environmental Protection Agency
EPT	Ephemeroptera, Plecoptera, and Trichoptera
FtHl	Foothills
GIS	Geographic Information System
HBI	Hilsenhoff Biotic Index
HQCW	high-quality cold water
km	kilometer
LOWESS	locally-weighted regression line
LRBS	Log Relative Bed Stability
LRBSfin	Log Relative Bed Stability final version (Kaufmann et al. 2008)
LRBS_NOR	LRBS without bedrock or hardpan
MCW	Marginal Cold-water
Mtn	Mountain
MWW	Marginal Warm-water
mg/L	milligrams per liter
n/a	Not applicable
NHD	National Hydrography Dataset
NMAC	New Mexico Administrative Code
NMED	New Mexico Environment Department
NMMSCI	New Mexico Macroinvertebrate Stream Condition Index
ntu	Nephalometric Turbidity Units
PCA	Principal components analysis
RBP	Rapid Bioassessment Protocol
RBS	Relative Bed Stability
Ref	Reference
RIVPACS	River Invertebrate Prediction and Classification System
RMSE	Root mean squared error
RW	Reach-wide
SABS	suspended and bedded sediment
SWIMS	Surface Water Information Management System

TMDL	Total Maximum Daily Load
TSS	Total Suspended Solids
U.S. EPA	United States Environmental Protection Agency
WQS	Water Quality Standards
WSA	Wadeable Streams Assessment
WW	Warm-water
Xer	Xeric

1 Introduction

The biological effects of excess fine suspended and bedded sediments in streams and rivers are well established in the scientific literature (e.g., Waters 1995, Wood and Armitage 1997, Suttle et al. 2004, Opperman et al. 2005). These effects include displacement of interstitial habitat space, clogging of water movement through sediments, disruption of normal predator-prey relationships through visual impairment, decreased primary productivity, increased macroinvertebrate drift, abrasion or smothering of gills and other organs, and increased uptake of sediment-bound toxicants. While these effects are well known, the process for establishing thresholds of sediment effects is rather new and evolving (U.S. EPA 2006, Paul et al. 2008, Cormier et al. 2008, Jessup 2009a, Bryce et al. 2008, 2010).

Sediments cannot be treated as introduced pollutants such as pesticides because they are not uniquely generated through human input or disturbance. Rather, sediments are components of natural systems that are present even in pristine settings and to which stream organisms have evolved and adapted. Therefore, the detection of a sediment imbalance is more difficult than detecting an absolute concentration or percentage that represents a clear biological impact.

The State of New Mexico Environment Department (NMED) associates imbalanced suspended and bedded sediment supply with effects to aquatic life uses through narrative or comparative standards (New Mexico Administrative Code 20.6.4.13). The degrees to which certain sediment quantities are unnatural and detrimental are as yet loosely defined. The U.S EPA Region 6 and NMED are interested in the development of benchmarks or thresholds, generally following the steps provided in U.S. EPA's Framework for developing suspended and bedded sediment (SABS) water quality criteria (U.S. EPA 2006), to better implement the existing narrative criteria. A workgroup of scientists from EPA, NMED, and Tetra Tech was convened to plan, organize, perform, and interpret analyses towards that end.

1.1 Problem statement

The condition of substrates in streams and rivers forms the foundation of habitat suitability for benthic organisms and their predators. The stability, size, shape, density, and porosity of the substrates directly affect the behavior of organisms and their ability to find food resources, hide from predation or other threats, and reproduce (Waters 1995, Wood and Armitage 1997). When substrate characteristics are out of balance in the channel or watershed, detrimental changes in the biological community structure and function can occur. Typically, increased disturbance leads to increased fine and mobile sediments, which leads to decreased stream habitat suitability.

Substrate and suspended sediment characteristics can be considered impacted at a site under two circumstances. First, they can be considered impacted if they are not similar to expectations for undisturbed sites in the same environmental setting. A second case for impact can be made when the substrate characteristics are detectably affecting the biota. In the first case, substrates may be more fine, more coarse, more unstable, or more stable than they are expected to be under broadly-recognized, undisturbed conditions (reference or best available conditions) for that particular environmental setting. This, in itself, can be an indication that streambed substrates are impacted by human disturbance. Biotic responses to disturbed substrates can be variable, but sub-optimal biotic conditions are often associated with unbalanced sediments. These relationships can be demonstrated through analyses. Both of these assessment strategies can be used to develop numeric benchmarks.

The purpose of these analyses is to identify sediment characteristics that are expected under the range of environmental settings in New Mexico, especially in undisturbed reference streams. Through this characterization, it will be possible to identify situations where the expectations are not met, using sediment indicators that show responsiveness to disturbance. Associating biological measures with sediment indicators will further indicate situations where the disturbance causes sediment imbalance and biologically-relevant habitat degradation. The results of these analyses include a set of recommendations regarding the application of quantitative sediment benchmarks on New Mexico perennial streams. This includes a range of possible benchmarks, as well as the rationale and the strengths and weaknesses of each. NMED will consider the recommendations to select benchmarks for proposed revisions to New Mexico's sedimentation assessment methodology. This will allow NMED to identify perennial stream reaches impacted by sediment for Clean Water Act 303(d) listing, to assist in the development and implementation of Total Maximum Daily Load (TMDL) targets to address sediment as a pollutant, and to potentially assist in the determination of successful watershed restoration activities.

Additional analyses were included to address questions related to future sampling and tiered assessments. By some estimates (Shuryn 2007), the expense of collecting the set of habitat variables typically collected by NMED is less than half as expensive as the more detailed set specified by the Environmental Monitoring and Assessment Program (EMAP). While the more detailed measures may allow more site-specific indictor measurements, the effort is substantial and may limit the number of sites monitored. NMED wanted a basic evaluation of the added value of the more detailed measures in terms of assessment accuracy, responsiveness to the stressor gradient, and relationships with biological conditions. Tiered assessments may use simpler measures as screening tools and more detailed measures for detailed assessments.

1.2 General Procedures

The approach to setting sediment benchmarks in New Mexico followed seven basic steps, as follows:

1. Review background information
2. Assemble datasets
3. Establish reference sites
4. Classify sites
5. Characterize sediments
6. Describe stressor-response relationships
7. Recommend benchmarks

These steps are loosely based on the U.S. EPA Framework for developing SABS water quality criteria (U.S. EPA 2006). The steps are also the basis for the organization of this report. We review the background requirements in the next section of this introduction, and then report the analytical methods and results of all other steps in the analysis chapter. A final section provides a synthesis discussion of the significance and limitations of the results, recommendations regarding sediment benchmark selection, critical data needs, and potential future refinements.

The analytical approach was to corroborate sediment expectations and sediment-biological relationships through multiple techniques. Three analytical techniques were used to identify potential benchmarks for sediments in New Mexico. These include analyses of reference distributions, quantile regression, and change-point analysis. Reference distributions describe expectations in least disturbed sites, which are classified by natural site types. Quantile regression and change-point analysis compare sediment conditions with biological conditions, using the biological conditions to indicate the degree to which aquatic life uses are supported. This report presents the results of each analysis and recommends possible benchmarks for NMED to consider. Selection of final sediment benchmarks using the weight-of-evidence from multiple analytical approaches will provide a solid basis for protective management strategies.

1.3 Background Information

New Mexico has narrative standards for both suspended and bedded sediments as well as a protocol for applying those standards. For bedded sediments, NMED hopes to translate existing narrative water quality standards related to sediment impacts into numeric measurements based on site class that can be used for assessment of stream conditions and suitability for aquatic life uses. The current NMED protocols for assessing bedded sediments include three steps (NMED 2009).

1) Directly evaluate instream habitat by measuring the amount of fine particles (defined in NMAC 20.6.4.13 as 2 mm or less) in representative riffles in both the site of concern and an appropriate best available reference site,
2) Compare differences, if any, between instream habitat of the site of concern and the best available reference site, and
3) Verify or confirm results obtained in number 2 by comparing benthic macroinvertebrate communities at the site of concern to the best available reference site or defined reference condition.

Perennial sites in any ecoregion in NM with less than 20% fines in a representative riffle are currently assumed to be meeting the narrative sedimentation criteria, regardless of comparison to a reference site or condition. Sites with greater than 20% fines are compared to a reference site or condition. The limitations of this approach lie in the reliance on data from one specific reference site, which may or may not be appropriate for comparison. Also, the 20% fines cutoff for all perennial streams does not take into account New Mexico’s varied ecoregions and associated geological and physiographic characteristics around the state. This protocol was developed to support an interpretation of the *State of New Mexico Standards for Interstate and Intrastate Surface Waters* narrative standard for bottom deposits found at NMAC 20.6.4.13 (NMWQCC 2007):

> A. Bottom Deposits and Suspended or Settleable Solids:
>
> (1) Surface waters of the state shall be free of water contaminants including fine sediment particles (less than two millimeters in diameter), precipitates or organic or inorganic solids from other than natural causes that have settled to form layers on or fill the interstices of the natural or dominant substrate in quantities that damage or impair the normal growth, function, or reproduction of aquatic life or significantly alter the physical or chemical properties of the bottom.
>
> (2) Suspended or settleable solids from other than natural causes shall not be present in surface waters of the state in quantities that damage or impair the normal growth, function or reproduction of aquatic life or adversely affect other designated uses.

At the outset of this project, we focused on three bedded sediment indicators to characterize the sediment conditions in New Mexico streams. These included the Relative Bed Stability (RBS) index (Kaufmann et al. 2008, 2009), and the areal percent of fines (<0.06 mm), and sand & fines (<2 mm) in the streambed. Data on these three indicators were relatively accessible for analysis because they have been collected by NMED, EPA, and agencies of neighboring states. In some sites, NMED collected bedded sediment data from representative riffle portions of the stream only, whereas other techniques sample the entire reach, regardless of habitat type. The data types were distinctly identified and analyzed.

For this document, percent sand & fines includes particles less than 2 mm. In the NM *Standards for Interstate and Intrastate Surface Waters*, this size fraction is called "% fines". This metric must be addressed in order to translate the current narrative standards into numeric values. Note that this document uses "% fines" to denote the fraction of particles less than 0.06 mm diameter.

For suspended sediments, more specific criteria address the measurement of turbidity, also found at NMAC 20.6.4.13 (NMWQCC 2007):

> J. Turbidity:
>
> Turbidity attributable to other than natural causes shall not reduce light transmission to the point that the normal growth, function or reproduction of aquatic life is impaired or that will cause substantial visible contrast with the natural appearance of the water. Turbidity shall not exceed 10 ntu over background turbidity when the background turbidity is 50 ntu or less, or increase more than 20 percent when the background turbidity is more than 50 ntu. Background turbidity shall be measured at a point immediately upstream of the turbidity-causing activity. However, limited-duration activities necessary to accommodate dredging, construction or other similar activities and that cause the criterion to be exceeded may be authorized provided all practicable turbidity control techniques have been applied and all appropriate permits and approvals have been obtained.

Prior to New Mexico's 2005 triennial review of water quality standards, numeric turbidity criteria existed for certain classes of waters. Specific numeric criteria generally only existed for streams of certain designated uses; those with expectations for clear water such as high quality coldwater and coldwater designated uses. Numeric turbidity criteria for sites designated for warm water and marginal warm water aquatic life uses did not exist, presumably because water clarity is not expected to be high even in natural waters of this class. New Mexico developed an interim turbidity assessment protocol for the 2010 Clean Water Act 303(d) listing cycle that utilized the previous numeric criteria as numeric translators. Both turbidity and Total Suspended Solids (TSS) have been collected with some regularity in New Mexico. Flow conditions at the time of collection of the suspended sediment sample can have substantial effect on the measurement. Therefore, when possible, samples with similar flow conditions were analyzed in groups. Low-flow conditions were most commonly sampled.

2 Analysis

As stated above, the analysis of suspended and bedded sediments in New Mexico will follow a series of steps, repeated here. Because the background information was presented in the introduction, the following sections begin with the second step.

1. Review background information
2. Assemble data sets
3. Establish reference sites
4. Classify sites
5. Characterize sediments
6. Describe stressor-response relationships
7. Recommend benchmarks

2.1 Assemble Data Sets

Data sources considered for analysis of sediment effects in New Mexico included the Ecological Data Application System (EDAS) maintained by NMED and multiple sediment datasets compiled by NMED, EPA, and neighboring states (**Table 1**). Data from states neighboring New Mexico were used if sites were in an ecoregion that was contiguous with those in New Mexico and the site was within 50 -150 miles of the state border (generally within 50 but extending to 150 in the east and northeast or as needed to include sites in the less densely sampled plains and xeric ecoregions). EMAP and Wadeable Streams Assessment (WSA) data were downloaded from an EPA-sponsored website. Data from NMED and neighboring states were delivered to Tetra Tech as databases or spreadsheets. Data were compiled in two databases, one for biological samples and one for sediment and ancillary data. From these databases, data were further formatted to work with analytical software.

Table 1. Datasets used in analyses.

Data set	Sampling Years	Source
NMED EDAS benthic macroinvertebrate data	1980 – 2007	NMED
NMED targeted riffle sediment data (taken during EMAP sampling)	2006 – 2007	NMED
EMAP West	1999 – 2004	EPA; SWIMS web site[a]
EMAP Wadeable Streams Assessment	2000 – 2005	EPA; SWIMS web site[a]
EMAP Arizona Streams	2007	EPA; SWIMS web site[a]
EMAP New Mexico	1999 – 2001, 2006 – 2007	EPA; SWIMS web site[a]
EMAP Region 8 Colorado Streams	1994 – 1995	EPA; SWIMS web site[a]
NMED suspended sediment data	2001 – 2009	NMED
Site GIS characterization	2000 – 2010	EPA Region 6 analysis

a: The Surface Water Information Management System (SWIMS) was maintained by the EPA National Health and Environmental Effects Research Laboratory-Western Ecology Division

Almost all sites had complete datasets, including 1) site characterization data (landscape measures and EMAP-style site, habitat, and catchment characterization), 2) sediment indicator variables, and 3) biological response variables. Geographic information system (GIS) data from previous analyses were retrieved from Tetra Tech archives, but found to be incomplete. Therefore, a GIS analysis was carried out by EPA Region 6 to consistently characterize land use, geology, and climatic conditions at each site and in the site catchments (**Appendix A**). Human disturbance variables analyzed at each site included land use, land cover, dams, road density, and road – stream intersections. The natural characteristics of sites and catchments included catchment area, National Hydrography Dataset (NHD) stream reach slope, land slope in the catchment and at the site, ecoregion designation, stream order, site elevation, soil permeability, precipitation, and geologic type (for erodibility ratings).

Geologic types identified in the GIS analysis were categorized for erodibility (or fine sediment production potential). Catchment characteristics in relation to geologic erodibility were determined by the percentage of the site catchment with geological types that are highly or moderately erodible (**Appendix B**). Stream power, the product of discharge and channel slope, is a measure of the capacity of a waterway to carry bedload. An index of stream power was calculated as the product of catchment area, stream slope, and precipitation.

Data were manipulated to refine, transform, average, or calculate metrics on sediment indicators, biological indicators, and site/catchment characteristics. Biological metrics and indices were calculated from the taxa lists. These included the New Mexico Macroinvertebrate Stream Condition Index (NMMSCI - Jacobi et al. 2006) and other metrics that were components of the NMMSCI or otherwise believed to be generally responsive to sediment stresses in New Mexico streams. The NMMSCI was standardized as a proportion of the reported impairment threshold in each of the biological site classes. These were 43.55 in the Low-Small sites, 51.64 in Low-Large sites, and 56.70 in High-Small sites. Because the NMMSCI was originally calibrated with midges (Diptera: Chirnomidae) at the family level, the calculations for this analysis also included family-level midge data. All other metrics counted midges and all other taxonomic groups at the level identified by the laboratory taxonomist (mostly genus-level). If benthic macroinvertebrate data existed from several sampling events, the metrics from the sampling event that coincided with the sediment sample were used. Where both targeted riffle and reach-wide sediment samples were collected, they were only compared to benthic samples collected from the same habitat type.

Sediment indicators that were of primary interest included RBS, % fines, and percent sand & fines for bedded sediment measures, and turbidity and TSS for suspended measures (**Table 2**). Other derivations of the sediment indicators were calculated from these basic measures. Such derivations included statistical transformations or standardization of the % fines and percent sand & fines to natural variables (i.e., determination of residuals). The RBS index was calculated from variables in the EMAP habitat files. If sediment data existed from several sampling events, these data were summarized for each site by averaging multiple sampling dates or replicates.

Table 2. Sediment indicators.

Sediment Indicator	Description
Relative Bed Stability	A measure of the relationship of the median particle size in a stream reach compared to the critical particle size calculated to be mobilized by standardized fluvial stresses in the reach. Median particle size is determined using a reach-wide pebble count (Peck et al. 2006). Critical particle size is calculated from channel dimensions, flow characteristics, and channel roughness factors (Kaufmann et al. 2008). The measure is expressed as a logarithm of the ratio of geometric mean to critical particle size and is abbreviated as “LRBS”.
Percent Fines	The percentage of systematically selected (with random start) streambed substrate particles that are ≤0.06 mm (silt, clay, or muck; not gritty when rubbed between fingers). Random particles were selected using either a reach-wide or targeted riffle pebble count. Abbreviated “%Fines”.
Percent Sand & Fines	The percentage of systematically selected (with random start) streambed substrate particles that are ≤2.0 mm in diameter. Random particles were selected using either a reach-wide or targeted riffle pebble count. Abbreviated “%SaFn”.
Turbidity	A measure of the degree to which light rays shining through the water are scattered by solid particles or water color. Measured in nephelometric turbidity units (ntu). Abbreviated “Turb”.
Total Suspended Solids	A measure of the amount of solids in water that are not in solution and that can be removed by filtration. Measured in mg/L. Abbreviated “TSS”.

2.2 Identifying reference (and stressed) sites

Indicators of the sediment and biological condition at a site were expected to correlate to the intensity of disturbance at the site and in the contributing watershed. Sites with minimal evidence of disturbance at the reach scale and in the catchments are expected to exemplify our best expectations for sediment and biological conditions (Stoddard et al. 2006). Conditions at these sites are the reference conditions. They are standards for comparison to other sites. Reference sites are also typically examined for patterns of natural variability that can be used to define site classes.

Site data were more complete for bedded sediment sites compared to suspended sediment sites. Therefore, reference designations for bedded sediment sites were based on the complete site data and previous reference designations, whereas reference for suspended sediment sites depended on attainment of designated uses, designations assigned in bedded analysis, and professional judgment. For bedded sediment sites, preliminary indicators of site reference status included existing designations or procedures used to designate reference sites in previous studies. These studies include EMAP (Stoddard et al. 2005), NMED benthic multi-metric index development (Jacobi et al. 2006), NMED benthic predictive model development (Paul 2008), and Colorado Department of Environment and Public Health (CDPHE) multimetric index development (Jessup 2009b). The preliminary reference designations were further scrutinized by selective reapplication of existing site criteria, application of new criteria for land use and road density, checking agreement of multiple reference designations, and evaluation of aerial imagery. Where designations were contradictory (one study indicates reference and another indicates degradation), then sites were left with no designation, termed “other”. Placing sites in this “other” category reduced the number of reference or stressed sites for subsequent analyses. However, this approach gave more credibility to the reference and stressed conditions and the benchmarks that were based on them.

The EMAP criteria for reference and stressed sites were modified to exclude measures of suspended and bedded sediment concentrations and then re-applied to all sites with supporting data. Excluding measures of sediments allows assessments of site quality that are independent of the stressor for which we are developing benchmarks. This avoids the circular reasoning that would result if we assumed that sites with low percentages of fine sediments were reference quality. Criteria included chloride, total phosphorus, total nitrogen, and three measures of riparian disturbance: all disturbances, agricultural disturbances, and crop related disturbances (**Table 3**) (Kaufmann et al. *in prep.*). Road densities statistics from EMAP were incomplete and were not reapplied in this set of criteria, though road statistics were regenerated and applied as part of the GIS analysis (see below). Application of the criteria was such that a positive point was tallied for a reference indication and a negative point was tallied for a stressed indication. Sites were potential reference or stressed sites if the total points were ≥ 3 or < 0, respectively.

Table 3. Anthropogenic disturbance screening criteria used by the National Wadeable Streams Assessment to characterize least-disturbed (“reference”), moderately-disturbed, and most-disturbed (stressed) stream reach sample sites. Observed values less than the first number (before /) indicate reference conditions for that variable. Values greater than the second value indicate stress.

Region E	coregion	Chloride (mg/L)	Total Phosphorus (mg/L)	Total Nitrogen (mg/L)	Riparian Disturb. (All)	Riparian Disturb. (Ag)	Riparian Disturb. (Crop)
MT-SW	23	300/1000	50/200	750/1000	1.25/2.5	0.25/0.6	0.2/0.5
MT-So. Rockies	21	200/1000	25/200	750/1000	1.25/2.5	1.0/1.4	0.2/0.5
PL-No. Cultivated	25	1000/2750	200/900	2000/4000	--	0.6/1.4	0.15/0.25
PL-Range	26	1000/3000	200/900	1000/3000	--	0.6/1.4	0.15/0.25
Xeric	20, 22, 24, 79	1000/2500	50/300	1000/4000	1.25/2.5	0.6/1.4	0.15/0.25

Reference sites used in the NM predictive biological model (RIVPACS, Paul 2008) were identified using professional judgment in combination with criteria specified in the 2006 multimetric index analysis (Jacobi et al. 2006). The Jacobi criteria included a Rapid Bioassessment Protocol (RBP) habitat score as one site evaluation factor among 10 factors. This habitat score included three sediment-related ratings of ten total ratings. Because the evaluation of stream sediments made up such a small fraction of the reference site selection factors (~3% of the factors), the NMED reference site designations were appropriate for development of sediment benchmarks.

Reference sites in Colorado were selected using criteria that purposefully avoided evaluation of habitat features, including any sediment assessments (Jessup 2009b). Therefore, Colorado reference site designations are useful for sediment benchmark development. Arizona reference site criteria require that a reference site will have habitat scores greater than 14 (out of 20) (ADEQ 2006). Twelve (12) points of the habitat evaluation are related to sediment deposition, embeddedness, and bank stability. Therefore, alternative reference criteria (EMAP and GIS) were used for sites in Arizona in lieu of the reference site designations of the Arizona Department of Environmental Quality (ADEQ).

GIS data on land uses, road densities in the site catchments, and nearby dams were also used to identify reference and stressed sites. Criteria for GIS variables (land uses, road density, and dam density) were established as in **Table 4.** Each variable was given a level at which disturbance was likely or not likely, based on distributions of data within the entire data set and professional knowledge of reasonable levels of impact. We had a goal that 10-20% of sites would be identified as either reference or stressed using the criteria. As an example, if the percentage of natural land cover in the catchment was greater than 99%, then the site was a candidate for a reference designation. The cut-off values were used as screening tools, since all sites with extremes of land use or road densities were reviewed through remote aerial imagery.

Table 4. Reference and stressed criteria for GIS variables.

Variable Reference	**threshold**	**Stressed threshold**
Natural land uses	> 99%	< 90%
Road density	< 0.2km/km^2	> 0.7km/km^2
Road crossing density	< 0.1/km^2	> 1.5/km^2
Dam density	< 0.05 dams/km^2	> 0.05 dams/km^2

We reviewed all potential reference and stressed sites with aerial imagery (GoogleEarth, available at: http://earth.google.com/) to identify any gross misclassifications that were evident from land use patterns. If any site showed potential for reference or stressed site status, then the site was examined in GoogleEarth to confirm the preliminary indication. In GoogleEarth, we checked for near-site conditions, presence/absence of stressors (roads, development, agricultural, mining, channel alteration, etc.). This review was used as a confirmation step, where reference or stressed site could be removed from the list, but no new sites were added through this review. NMED confirmed or rejected sites based on mapped information, site familiarity, and professional judgment, leading to the final lists of reference and stressed sites.

Sites that showed ambivalent or contradictory reference characteristics and uncertain imagery in GoogleEarth were called "Other". The “Other” category also includes sites with intermediate level of disturbance or insufficient data for categorization. The relative sensitivity of indicators was assessed over a gradient of disturbance – Reference, Other, and Stressed.

In the suspended sediment dataset, there were some sites in common with the bedded sediment sites. The reference designations established for bedded sediments were adopted for the common sites in the suspended sediment analysis. In addition, the NMED Aquatic Life Use (ALU) impairment designations were used to identify sites with conditions supporting their uses. Lacking the data needed to more definitively determine the appropriate category for all 692 sites with available suspended sediment data, it was assumed that sites with Full Support designations were similar to reference, or relatively undisturbed. It should be noted that a site with an aquatic life use of Full Support does not expressly indicate that data were available and assessed for every single applicable aquatic life criteria. NMED did not necessarily assess for turbidity and/or bedded sedimentation (or all associated ALU criteria for that matter) at all sites, which may result in inconsistent measures of site disturbance.

Results: Reference Site Identification

Bedded Sediment Analysis

The screening and criteria process resulted in identification of 99 reference sites of the 229 sites used for the bedded sediment analysis (**Figure 1**). Most of the reference sites were in mountainous ecoregions which include the Southern Rockies (N = 37) and the Arizona/New Mexico Mountains (N = 41). Far fewer (25) stressed sites were identified, which were distributed relatively evenly among the ecoregions. Though there were numeric reference criteria for some variables, designations were made with subjective input (weight-of-evidence) because of the combination of multiple indications (numeric criteria, previous designations, screening examination of aerial imagery, and professional judgment of familiar sites).

The thresholds used as reference criteria by EMAP, NMED, and CDPHE were already established and were not adjusted. Criteria for GIS variables (land uses, road density, and dam density) were established as in **Table 4**. When these criteria were exceeded, they were used to confirm other indicators of reference status.

Because sites may have been evaluated with multiple previous analyses, the following statistics (**Table 5**) overlap among programs, giving the appearance that more than 99 sites were designated as reference. The EMAP criteria that were adjusted to remove sediment variables resulted in 44 sites that met four of the six reference criteria and 56 that met three criteria. After comparing to other reference designations and examining maps, 82 of those sites were designated as reference. Twenty three sites failed one or more of the six EMAP stressed criteria, resulting in 18 stressed site designations after additional screening.

The RIVPACS analysis identified 24 reference sites in New Mexico, 21 of which were confirmed. Likewise, NMED designations in the EDAS database (used for biological indicator development) identified 24 reference sites, 20 of which were confirmed. Only one of five stressed sites identified in EDAS were confirmed. In Colorado, 15 of the 17 sites identified by CDPHE were confirmed as reference for the current analysis. Only one stressed site in Colorado was identified and confirmed. Five stressed sites were identified based on GIS information alone.

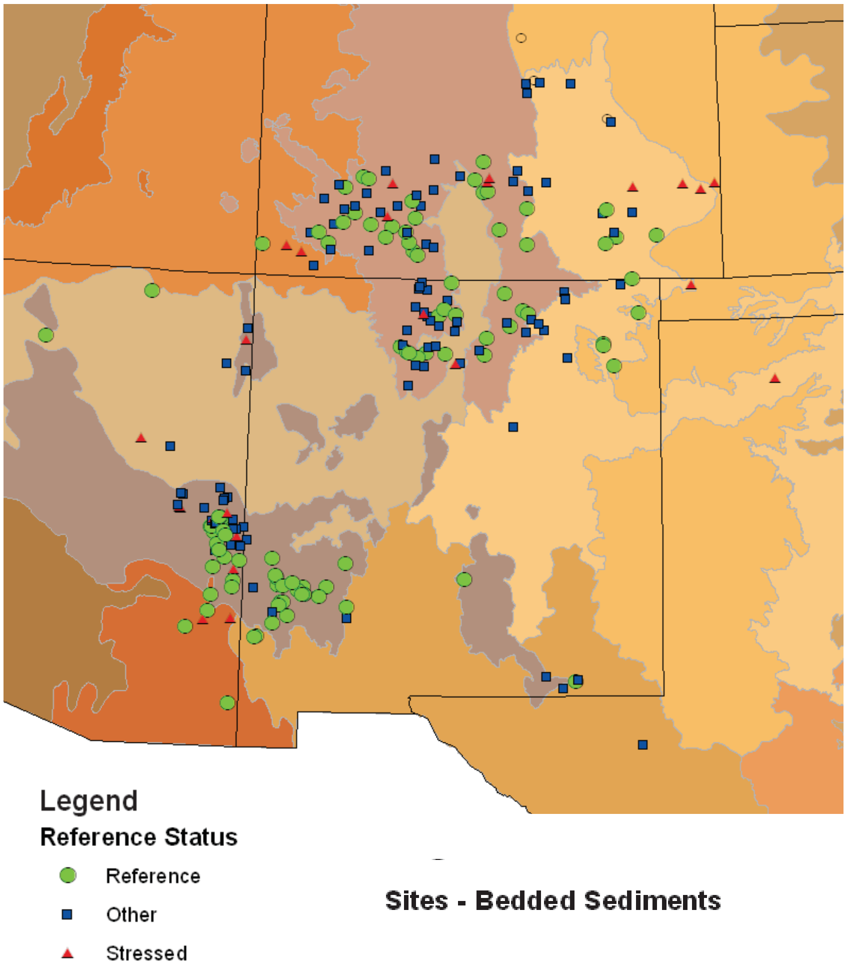

Figure 1. Site locations, reference status, and level 3 ecoregions of bedded sediment sites in New Mexico and surrounding areas.

Table 5. Reference and stressed sites indicated by individual criteria sets and as finally designated after review of consistent indications and aerial imagery.

Criteria Set	Potential Reference	Final Reference	Potential Stressed	Final Stressed
EMAP	100	82	23	18
RIVPACS	24	21	n/a	n/a
NMED EDAS	24	20	5	1
CDPHE	17	15	1	1
GIS	n/a	n/a	5	5

Results: Reference Site Identification
Suspended Sediment Analysis

Of the 692 sites with suspended sediment data, 186 were supporting their uses, which suggested relatively undisturbed conditions (**Figure 2**). An additional 48 sites were identified as reference because they were determined to be reference in the bedded sediment analysis. As with the bedded sediment sites, most of the sites with fewer stressors were in the mountainous ecoregions, the Southern Rockies (N = 93) and the Arizona/New Mexico Mountains (N = 43). While the lack of impairments was used to identify sites with relatively undisturbed conditions, the presence of impairments was not used to identify any more refined gradient of site disturbance. Therefore, sites were only categorized as relatively undisturbed and somewhat disturbed.

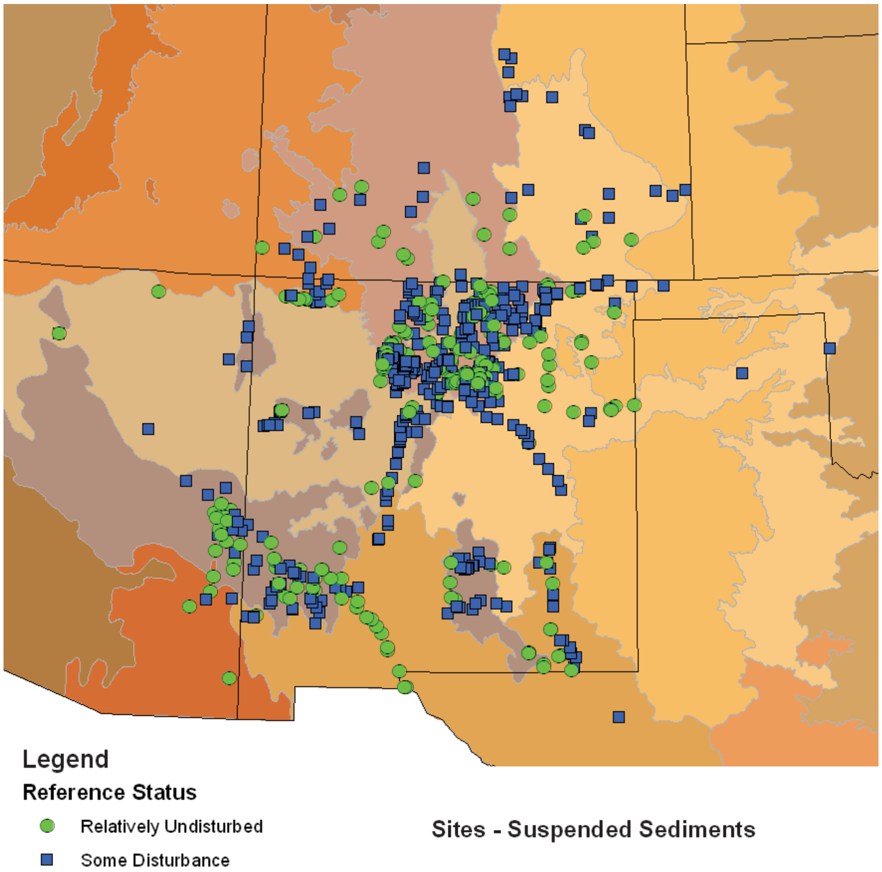

Figure 2. Site locations, reference status, and level 3 ecoregions of suspended sediment sites in New Mexico and surrounding areas.

2.3 Site Classification

Site classification is the process by which natural gradients among sites are examined to identify appropriate classes or “bins” of sites with similar sediment characteristics. The purpose of stream classification is to minimize within-class natural variability of sediment indicators so that anthropogenic disturbance can be recognized with less background noise. Potential site classification variables, sediment indicators, and biological variables were analyzed simultaneously to identify patterns of covariance. The analysis included several analytical methods (e.g., principal components analysis, correlation, and examination of bi-plots and distributions). Some additional techniques were exploratory and were not carried beyond preliminary stages, though they were sufficient to guide final analyses.

Ideally, reference sites alone would be used for site classification. However, we expect that the sediment indicators may respond differently to stress in the landscape depending on some natural characteristics, so it made sense to review patterns in all sites, not just reference sites. For instance, reference sites in erodible and resistant lithologies may have similar sediment signatures. However, with equal disturbance, the sites with erodible characteristics may have a much more profound signal of sediment stress, as was reported for Pacific Northwest streams (Kaufmann et al., 2009). We also hypothesized that ecoregions and stream size (or power) may be important determinants of natural sediment conditions. Aggregate ecoregions used in the EMAP-West study — Mountains, Plains, and Xeric — were considered as a starting point for stream classification, but were not accepted without scrutiny.

We started looking at classification options using the bedded sediment dataset because we had site data associated with each sample. The suspended sediment sites were examined using available classification data, including ALU categories and level 4 ecoregions. Classification options that were similar among bedded and suspended sediment indicators were emphasized so that the classification scheme in New Mexico would be simpler to apply among indicators.

Principal components analysis (PCA) was used as a primary tool for classifying sites. The analysis helps to structure the data based on related variance among variables so that the major patterns among variables can be interpreted. The variables can be entered into the model as primary determinants or as supplemental covariants. We included all sites in the analysis (not just reference sites) and included all natural, stressor, and indicator variables with continuous value distributions as the primary determinants. Biological variables were supplemental. Variables were transformed as needed to approximate a normal distributions using logarithmic and ArcSine-SquareRoot transformations.

Correlation analysis was used to describe basic relationships between sediment and environmental variables in reference sites. The Pearson product-moment correlation coefficient was calculated for a matrix of individual variables (sediment, natural, stressor, and benthic metrics) transformed as noted for the PCA analysis. The analysis was limited to reference sites when comparing individual pairs of variables to diminish some of the co-variability with stressors. Correlation results thus offer a slightly different perspective and new information compared to the PCA.

The relationships that were suggested by PCA and correlations were examined in box plots and bi-plots. Box and whisker plots were compared among expected site classes, such as ecoregions, to aid assessment of sediment variability. Ecoregions were examined using box plots because they are categorical and are one of the *a priori* variables of interest. Such illustrations of the distributions of values are simple to generate and easy to interpret for one variable at a time. They are not so useful for detecting interactive effects. Bi-plots were used to show patterns of relationships between two variables and to highlight tertiary attributes of the relationships such as reference status, ecoregion, or other covariants. LOWESS regressions were used to show trends at local portions of the gradients and to identify possible change-points (see Section 2.5).

Classification results

Sites were classified so that our expectations of sediment indicator values are calibrated to the many environmental settings in New Mexico. Preliminary analyses indicated that sediment indicators (LRBS, % fines, and % sand & fines) were most strongly related to environmental variables related to stream size, stream slope, precipitation, and riparian vegetation. The workgroup had an *a priori* assumption that sediment conditions would vary by ecoregion. This assumption was somewhat validated by the preliminary results, which identified potential classification variables that are integral to the ecoregional classification scheme. Sediment conditions among ecoregions were investigated using box plots because ecoregional categories could not be interpreted in the PCA and correlation analyses.

We put extra effort into investigating lithology and stream power as possible classification variables, based on previous experience in Oregon (Jessup 2009). Lithology describes the underlying geology at a site or in the site's catchment. The types of rocks in the study area were classified regarding their potential for erosion and subsequent contribution of fine sediments to the streams. Stream power, the product of discharge and slope, is a measure of a stream's capacity for transporting its sediment load. We calculated an index of stream power as the product of catchment area × channel slope × precipitation.

In the PCA, the first three factors explained 53% of the variability in the primary variables (**Table 6**). Sediment indicators were most strongly related to the first axis, and somewhat related to the second axis. The first principal axis was related to the natural variables, catchment area, elevation, stream slope, and precipitation and to the density of road crossings in the site catchment. Benthic macroinvertebrate metrics were also most strongly related to the first axis, and were therefore related to the sediment measures. The second axis was related to stream size (power and width). The third and fourth axes were related to site location (latitude and longitude) and land uses. The third and fourth axes were not strongly related to sediment or biological measures.

Correlation analysis identified relationships among variable pairs, particularly between the sediment indicators and the environmental variables in reference sites. Many of the variables that were prominent in the PCA also showed strong relationships in the correlation analysis, such as slope, precipitation, elevation, and catchment area. More variables were significantly related to % sand & fines than to % fines or LRBS (**Table 7**). In a similar pattern, residuals of the % fines and % sand and fines variables had weaker correlations than their unadjusted counterparts. Natural variability in bed sediment fines derived from factors at the local site scale is partially accounted for by use of the LRBS and residual percentage variables. LRBS scales bed texture by the shear stress available for transporting bed sediments, adjusting for features of channel roughness (wood and pools) that diminish the shear stress exerted on bed particles themselves. The LRBS and LRBS_NOR variables were most strongly related to riparian vegetation, riparian disturbance, and precipitation. Stream power was initially a variable of interest for classification. It did not discriminate in the correlation analysis although the components of stream power did (slope, area, and precipitation).

Bi-plots illustrate the relationships suggested in the PCA and correlation analysis. The relationship between % sand & fines and catchment size appears somewhat imprecise in the bi-plot, though there seems to be a difference in sediment conditions in sites with catchments <100 km^2 compared to larger catchments (**Figure 3A**). The relationships that compare % sand & fines to precipitation and slope seem clearer, with sites having lower precipitation (<25cm/year) and slopes (<2.5%) also having more fine sediments (**Figure 3B**).

Table 6. PCA factor scores on the most important variables in all sites. Scores with magnitude greater than 0.60 are shown in bold-type and considered to be strong relationships. For descriptions of the residual calculations, see Section 2.4.1.

Variable code	**Variable description**	**F. 1 30%**	**F. 2 12%**	**F. 3 11%**	**F. 4 6%**
LRBS_fin	Relative Bed Stability index (log10)	0.56	0.45	-0.17	-0.10
LRBS_NOR	RBS without bedrock or hardpan (log10)	0.58	0.44	-0.18	0.07
asPCT_SAFN	% sand & fine sediments at the site (arcsin(sqrt())	**-0.72**	-0.44	0.20	0.21
asPCT_FN	% fine sediments at the site (arcsin(sqrt())	-0.58	-0.48	0.20	0.12
Resid_pSAFN	% sand & fines (residual of critical diameter)	**-0.61**	-0.39	0.10	-0.13
Resid_pFN	% fines (residual of critical diameter)	-0.43	-0.42	0.10	-0.03
Res2pSAFN	% sand & fine (residual as per Stoddard)	-0.30	-0.38	0.30	-0.26
Res2pFines	% fines (residual as per Stoddard)	-0.15	-0.39	0.26	-0.12
LRdX_km2	Road crossings per km2 in the catchment	**-0.86**	0.37	0.10	-0.04
LArea_km2	Catchment area (log10(km2))	**-0.84**	0.47	-0.04	0.04
LSTREAMSLOP	Stream slope (log10(%), NHD data)	**0.84**	-0.08	-0.09	-0.06
ELEV_m	Site elevation (m)	**0.78**	-0.15	-0.20	0.22
STREAMORDE	Stream Order (Strahler)	**-0.75**	0.50	0.12	0.00
Precip	Precipitation (cm)	**0.72**	0.14	-0.24	0.08
LXSLOPE	Stream slope (log10(%), field data)	**0.66**	0.07	-0.33	-0.16
LPower	Stream power (log10(Precip*Area_km2*Xslope))	-0.50	**0.65**	-0.28	-0.03
LXWIDTH	Average site width (log10)	-0.46	**0.63**	-0.29	-0.14
Point_X	Latitude of sample	-0.36	-0.26	**-0.63**	0.10
Point_Y	Longitude of sample	0.00	-0.36	**-0.73**	0.10
as_Nindx	Natural land uses (%)	0.35	0.45	0.21	**0.72**
U_INDEX	Developed (urban) land uses (%)	-0.37	-0.45	-0.21	**-0.71**
*TotalTax	Total taxa (count)	0.42	0.20	0.04	-0.05
*BeckBI	Beck's Biotic Index (weighted count of sensitive taxa)	**0.63**	0.24	-0.19	-0.01
*IntolTax	Number of taxa intolerant of pollution (count)	**0.65**	0.17	-0.23	0.00
*EPTTax	EPT taxa (count)	0.46	0.39	-0.20	-0.09

Table 7. Correlations (Pearson r) of the environmental variables most strongly related to the sediment variables in reference sites. Sediment variable codes are as described in Table 6.

Variable code	Variable description	asPCT FN	Resid pFN	Res2 pFines	asPCT SAFN	Resid pSAFN	Res2 pSAFN	LRBS fin	LRBS NOR
Point_X	Latitude of sample	0.31	0.26	0.09	0.22	0.15	-0.10	-0.14	-0.19
LArea_km2	Catchment area (log10(km2))	0.25	-0.01	-0.23	0.37	0.19	-0.08	-0.11	-0.14
PFORnew	Forest in the catchment (%)	-0.38	-0.36	-0.05	-0.30	-0.29	0.11	0.23	0.31
pctModHiErod	Moderately and highly erodible rocks (%)	0.25	0.17	0.03	0.16	0.11	-0.08	-0.07	-0.10
STREAMORDE	Stream Order (Strahler)	0.17	-0.05	-0.25	0.31	0.17	-0.06	-0.09	-0.11
ELEV_m	Site elevation (m)	-0.36	-0.17	0.13	**-0.51**	-0.36	0.00	0.30	0.36
Precip	Precipitation (cm)	**-0.47**	-0.21	0.07	**-0.64**	**-0.44**	-0.10	**0.41**	**0.46**
LSTREAMSLOP	Stream slope (log10(%), NHD data)	**-0.47**	-0.12	0.03	**-0.57**	-0.27	-0.07	0.25	0.28
LXSLOPE	Stream slope (log10(%), field data)	**-0.58**	-0.15	-0.17	**-0.64**	-0.22	-0.19	0.20	0.26
LPERCENTSLO	Land slope (log10(%))	-0.35	-0.08	-0.05	-0.36	-0.15	-0.08	0.18	0.20
XCL	Riparian Canopy > 0.3m DBH	-0.28	-0.06	-0.02	-0.35	-0.15	-0.09	0.19	0.22
XCMGW	Rip Veg Canopy+Mid+Ground Woody Cover	**-0.42**	-0.29	-0.19	**-0.41**	-0.32	-0.17	**0.41**	**0.44**
LRdX_km2	Road crossings per km2 in the catchment	0.37	0.15	-0.14	**0.45**	0.30	-0.06	-0.28	-0.33
W1_HALL	Index of riparian disturbance	**0.41**	0.26	0.07	**0.49**	0.35	0.10	-0.34	**-0.40**

DBH = Diameter Breast Height

Rip Veg = Riparian Vegetation

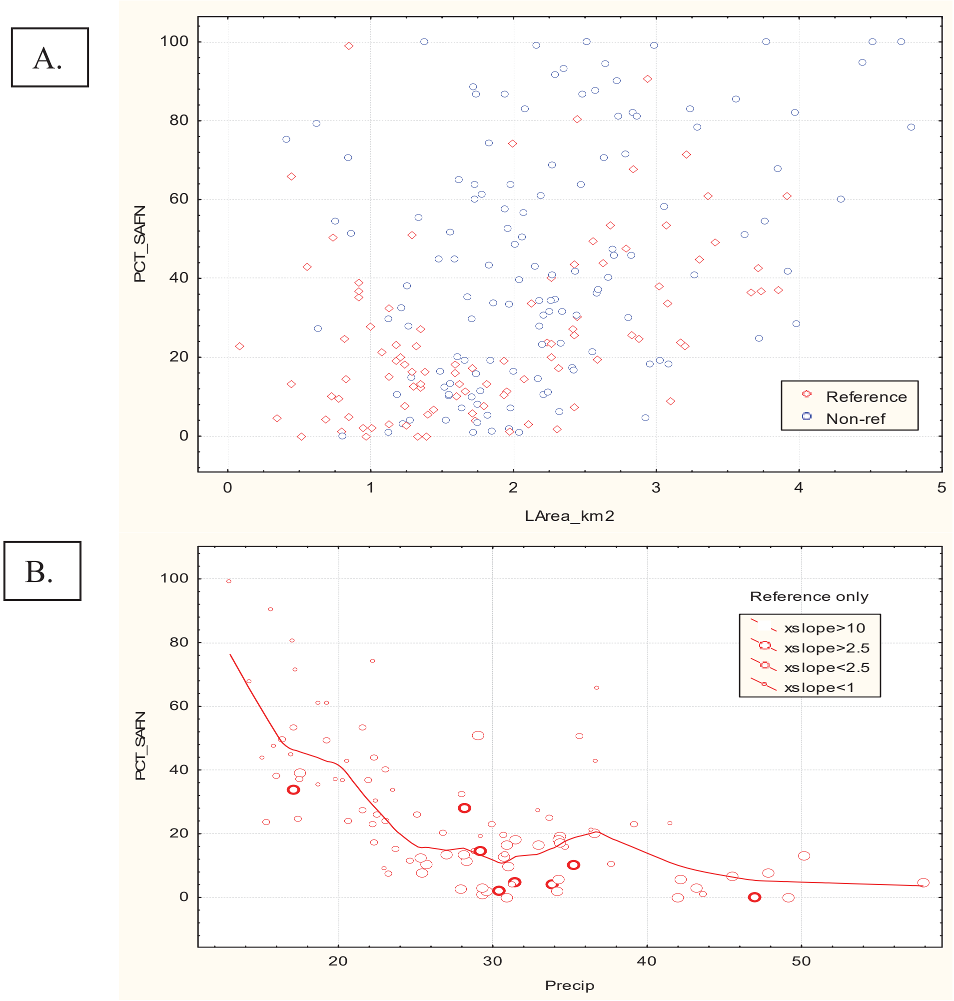

Figure 3. Relationships between % sand & fines and the influential variables A) catchment area and B) precipitation and slope. The LOWESS regression line is displayed in the plot of precipitation.

Similar relationships can be seen with % fines and slope (**Figure 4A**) and precipitation (**Figure 4B**). The cutoff points dividing higher and lower % fines are similar to the breakpoints described for % sand & fines (precipitation at 25 cm/year and slope at 2.5%). As can be seen in **Figure 4B**, precipitation varies in ecoregional categories.

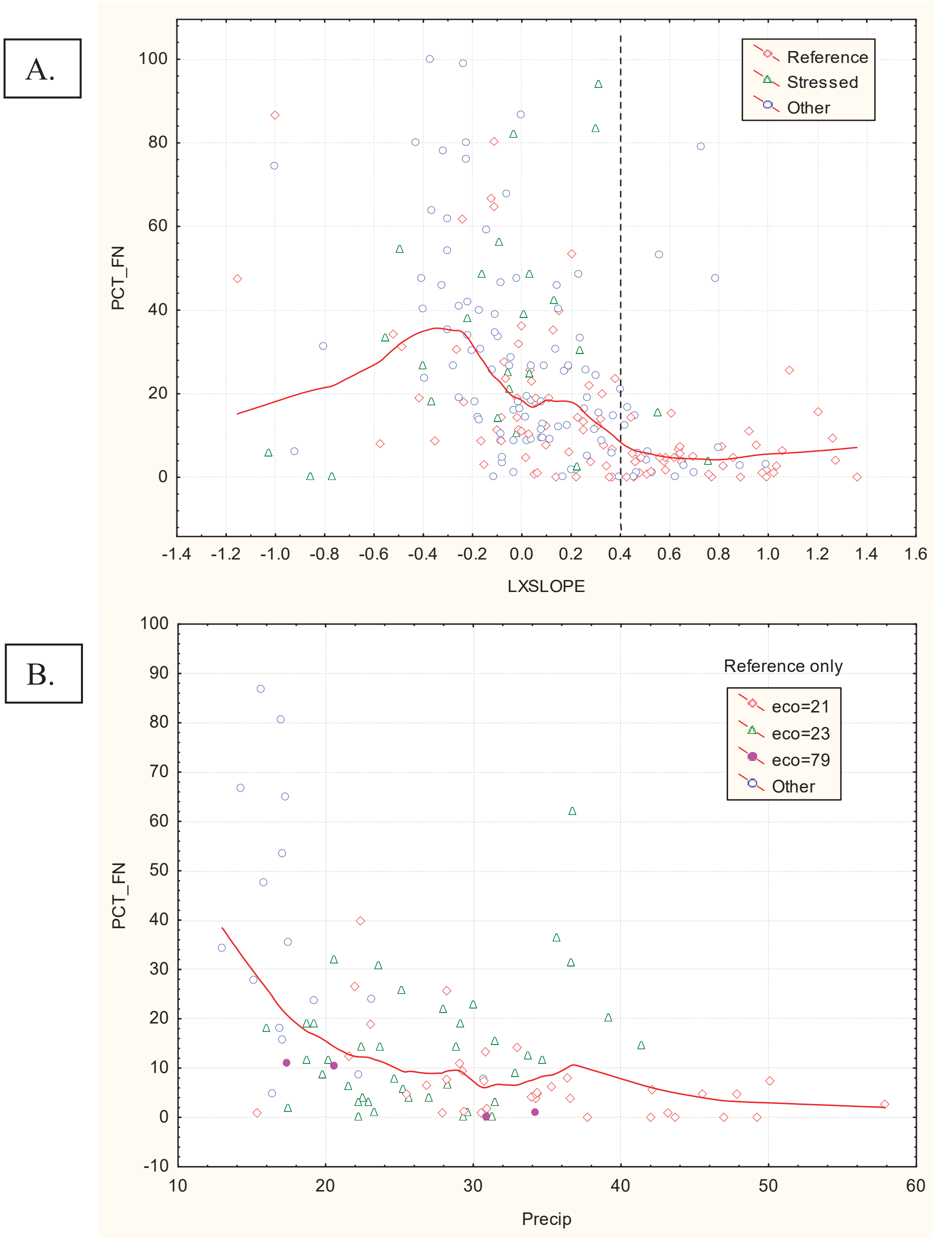

Figure 4. Relationships between % fines and the influential variables A) stream slope and B) precipitation and ecoregion. The LOWESS regression line is shown in both plots. The vertical dashed line describes a potential breakpoint at a slope of 2.5% (back-transformed from the log value).

With LRBS, a change-point can be seen at a precipitation level near 25 cm/year just as was observed with the other sediment indicators (**Figure 5A**). Though the relationship between LRBS and woody riparian vegetation showed the strongest correlation among reference sites (**Table 7**), the relationship in the bi-plot appears somewhat imprecise and a change-point would be hard to identify (**Figure 5B**). Woody riparian vegetation could vary in the landscape either naturally (some regions are natural grasslands or sparsely vegetated) or with human disturbance (clearing for development, silviculture, or agriculture).

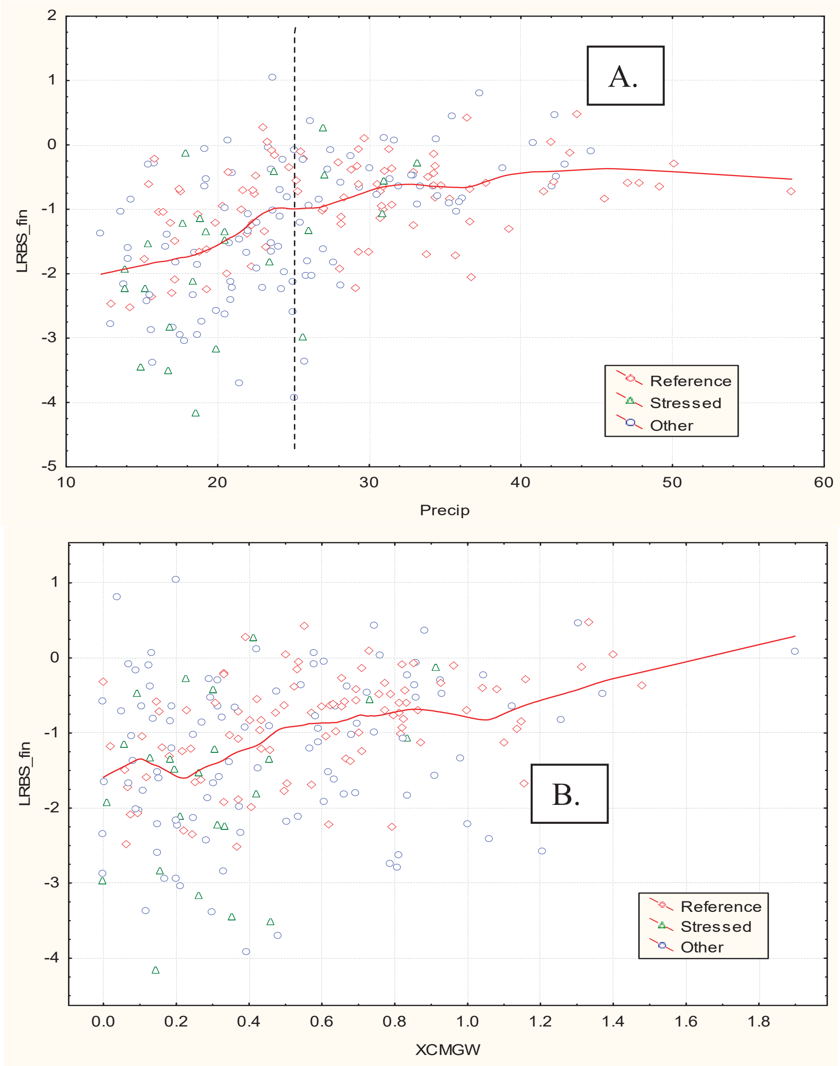

Figure 5. Relationships between LRBS and the influential variables A) precipitation and B) woody riparian vegetation (XCMGW). The LOWESS regression line is shown in both plots.

Because precipitation was an important determinant of all three sediment indicators and because ecoregions incorporate climatic conditions, we plotted precipitation and LRBS in the level 3 ecoregions (**Figure 6**). The ecoregions are effective predictors of precipitation, as well as of other influential determinants such as stream slope (not shown). Using ecoregions in a classification scheme was attractive because there are precedent classification schemes in the region (e.g., EMAP, Colorado bioassessments, New Mexico nutrient assessment protocol) and because an ecoregional scheme is relatively easy to conceptualize, communicate, and apply through mapping techniques. Classifying sites distinctly by ecoregion would be simpler than classifying by a continuous variable such as precipitation. Mountainous and non-mountainous ecoregions appear to be good classifiers of sediment conditions. They incorporate those underlying natural conditions that PCA and correlations showed to affect sediment conditions.

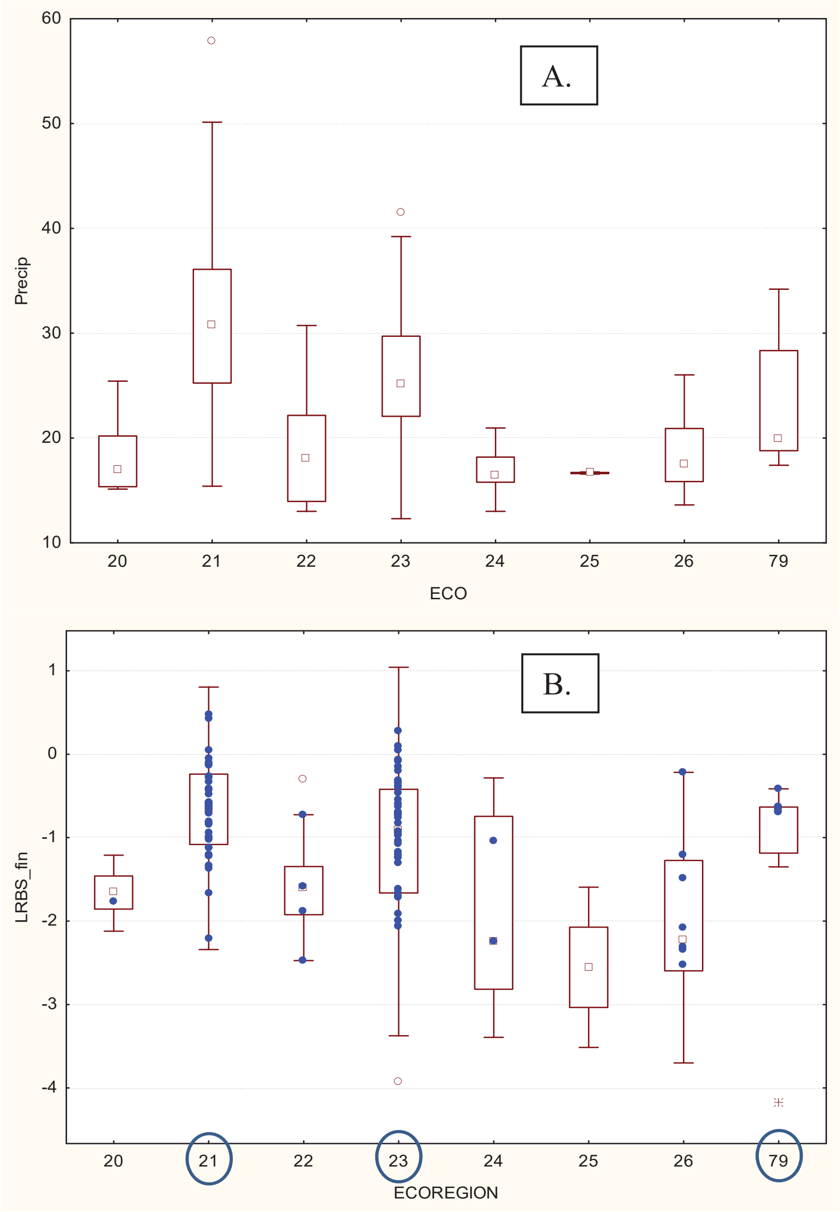

Figure 6. Precipitation (A) and LRBS (B) in the level 3 ecoregions. Ecoregions 21, 23, and 79 are generally mountainous compared to the others. The dot markers in B are data points for reference sites.

The level 3 ecoregions span considerable types of landforms in some cases. Therefore, level 4 ecoregions were considered for refinement of the classes. While a general division between mountains and plains was suggested (e.g., **Figure 6**), an additional transitional class seemed appropriate and necessary. Examination of sediment indicator values in level 4 ecoregions (**Figure 7**) and consideration of ecoregion descriptions for regions that were poorly represented in this dataset lead to site classification similar, yet modified from the EMAP mountains, plains, and xeric regions. For this analysis, the regions are termed Mountains, Foothills, and Xeric. Classes are defined as in **Table 8** and shown in **Figure 8**. This scheme recognizes the differences between high elevation, steep sloped, lush vegetation mountain streams; lower and drier foothill streams; and flatter and still drier xeric streams. In Arizona, level 4 ecoregions have not been defined. Therefore, assignment of sites to the Mountains or Foothills in ecoregion 23 in Arizona was based on review of aerial imagery with respect to vegetation, elevation, and apparent slope. The first factor scores of the PCA were significantly different (ANOVA followed by Tukey's Honestly Significant Difference test, p<0.0001) among the site classes (**Figure 9**), indicating that the site classes have somewhat distinct environmental characteristics.

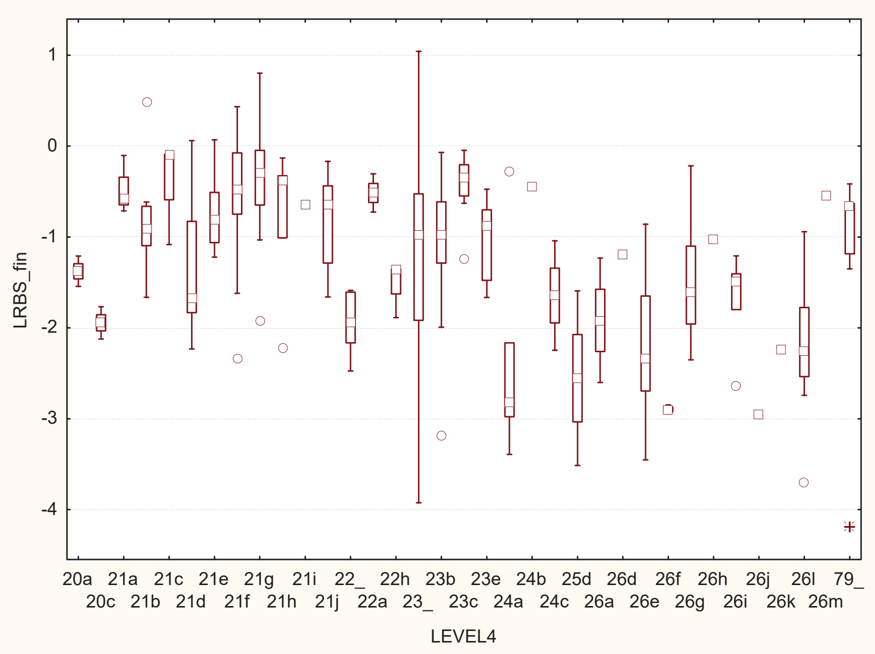

Figure 7. LRBS values in level 4 ecoregions. Level 4 ecoregions have not yet been defined in Arizona, resulting in level 3 designations for some sites in ecoregions 22, 23, and 79.

Table 8. Definition of bedded sediment site classes, including names of the level 3-4 ecoregions used to define the classes.

Site Class	Definition
Mountains	Ecoregions 21 and 23, except 21d, 23a, 23b and 23e
Foothills	Ecoregions 21d, 23a, 23b, 23e and 79
Xeric	Ecoregions 20, 22, 24, 25, and 26

Ecoregion number	Ecoregion Name
20	Colorado Plateaus
21	Southern Rockies
21d	Foothill Woodlands and Shrublands
22	Arizona/New Mexico Plateau
23	Arizona/New Mexico Mountains
23a	Chihuahuan Desert Slopes
23b	Madrean Lower Montane Woodlands
23e	Conifer Woodlands and Savannas
24	Chihuahuan Deserts
25	High Plains
26	Southwestern Tablelands
79	Madrean Archipelago

Figure 8. NM Mountain, Foothills, and Xeric site class map.

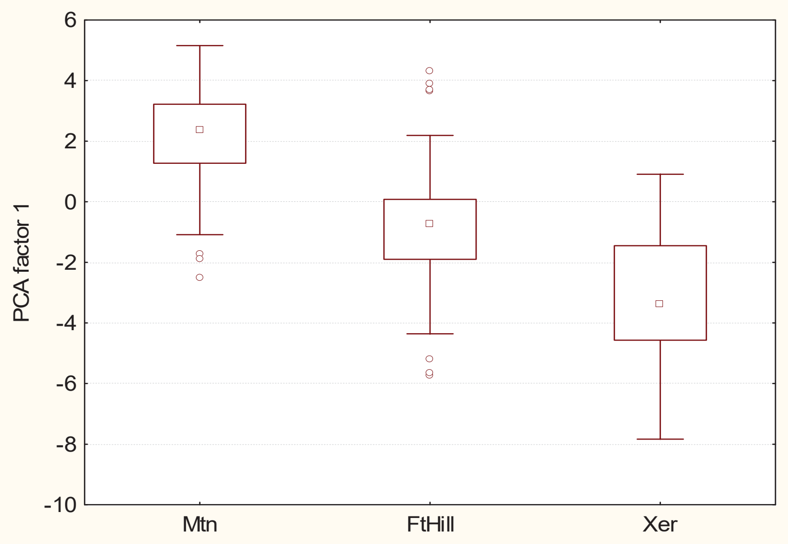

Figure 9. PCA scores in all bedded sediment sites.

Lithologic erodibility was not found to be an important determinant of sediment classes in the PCA and correlation analyses. However, the following plot (**Figure 10**) suggests that in sites with resistant lithology, sediment conditions are less responsive to disturbance than in erodible sites. In both the mountainous and non-mountainous regions, reference sites have relatively consistent conditions across the erodibility scale. Stressed sites do not show excessively low LRBS values until the percentage of moderately and highly erodible rock cover in the catchment is greater than 40%. This pattern was also noted for the other sediment indicators (not shown). Excessive sediments in streams are less likely in areas with resistant rocks because the sources of fine particles are less plentiful and few particles are mobilized even in response to disturbance. This might be a consideration in applications of any benchmarks resulting from these analyses.

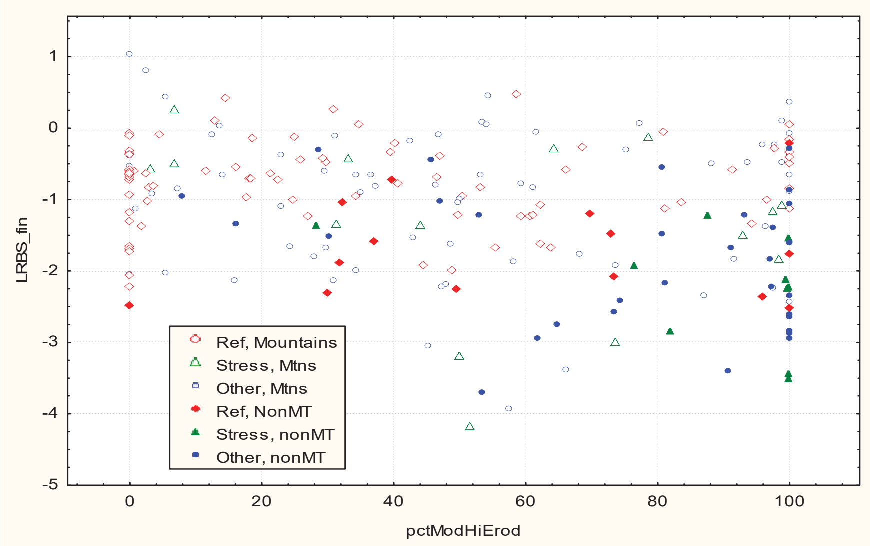

Figure 10. LRBS values in relation to the percentage of moderately and highly erodible lithology in the site catchments.

Classification results for suspended sediment sites

Because the site specific data were incomplete for sites with suspended sediment samples, there were few options for analyzing natural variability related to suspended sediments. There were two options for a priori classification: 1) based on ALU categories, and 2) based on classes developed for bedded sediment sites.

The ALU categories were available with the initial data set for all sites within New Mexico. Initial comparisons of suspended sediment measures used ALU categories because they were readily available. The five initial categories were reduced to three categories by combining two cold water types (MCW and CW) and combining two warm water types (MWW and WW). To determine a final classification scheme, subsequent analyses compared site assignments and sediment indicator value distributions among the two classification schemes (ALU designations and bedded sediment classes).

Most of the high-quality cold water (HQCW) sites were in the Mountain site class and most of the warm water (WW) sites were in the Xeric site class (**Table 9**). The Foothills had the fewest sites that were supporting their uses. These were mostly cold water (CW) sites, but in general the Foothills contained a mix of all ALU types. Because there were similarities in the site assignments among the two classification schemes, the suspended sediment value distributions were similar (**Figure 11**). In sites supporting their uses, the mean values were comparable among classification schemes (HQCW – Mtn; CW – FtHl; WW – Xer). Variability was lower in the bedded sediment classes, except in the case of turbidity in the CW – FtHl comparison. Bedded sediment classes were used as the final classification scheme because a.) a parallel classification scheme for the two indicator categories would simplify application and communication of any resulting benchmarks, b.) the bedded sediment measures were usually less variable using the bedded sediment scheme, and c.) sites outside of New Mexico could be assigned to bedded sediment classes, which is not so for ALU classes.

Table 9. Comparison of Aquatic Life Use classes and bedded sediment classes.

ALU \ Site Class	Mtn Supp	Mtn Other	FtHl Supp	FtHl Other	Xer Supp	Xer Other	Row Total
HQCW_supp	61		7		8		76
HQCW_other		99		57		21	177
CW_supp	9		12		24		45
CW_other		21		38		58	117
WW_supp	2		7		56		65
WW_other		31		23		133	187
All Grps	72	151	26	118	88	212	667

2.4 Indicator Characteristics

The sediment indicators currently in use and proposed for use in New Mexico streams were scrutinized to determine which were most appropriate for further analysis and application in the regulatory context. Sediment indicators should be selected based on: (a) association with designated uses; (b) availability and accessibility of data; (c) reliability of measurement characteristics; (d) appropriateness for the proposed analytical methods; and (e) applicability in the regulatory context. The short list of sediment indicators (**Table 2**) was reviewed to see if analyses should be limited or the variables should be adjusted to account for natural variability.

In this section we further describe the sediment indicator characteristics. For bedded sediments, we describe the adjustments to the absolute values that account for natural conditions, sediment conditions relative to stressors, and the statistics for bedded sediment concentrations in reference sites. For suspended sediments, we explore the relationships between turbidity and TSS, concentrations relative to increased flows, suspended sediments in relation to bedded sediments, statistics for suspended sediment concentrations in relatively undisturbed sites, and suspended sediment concentrations in relation to stressors.

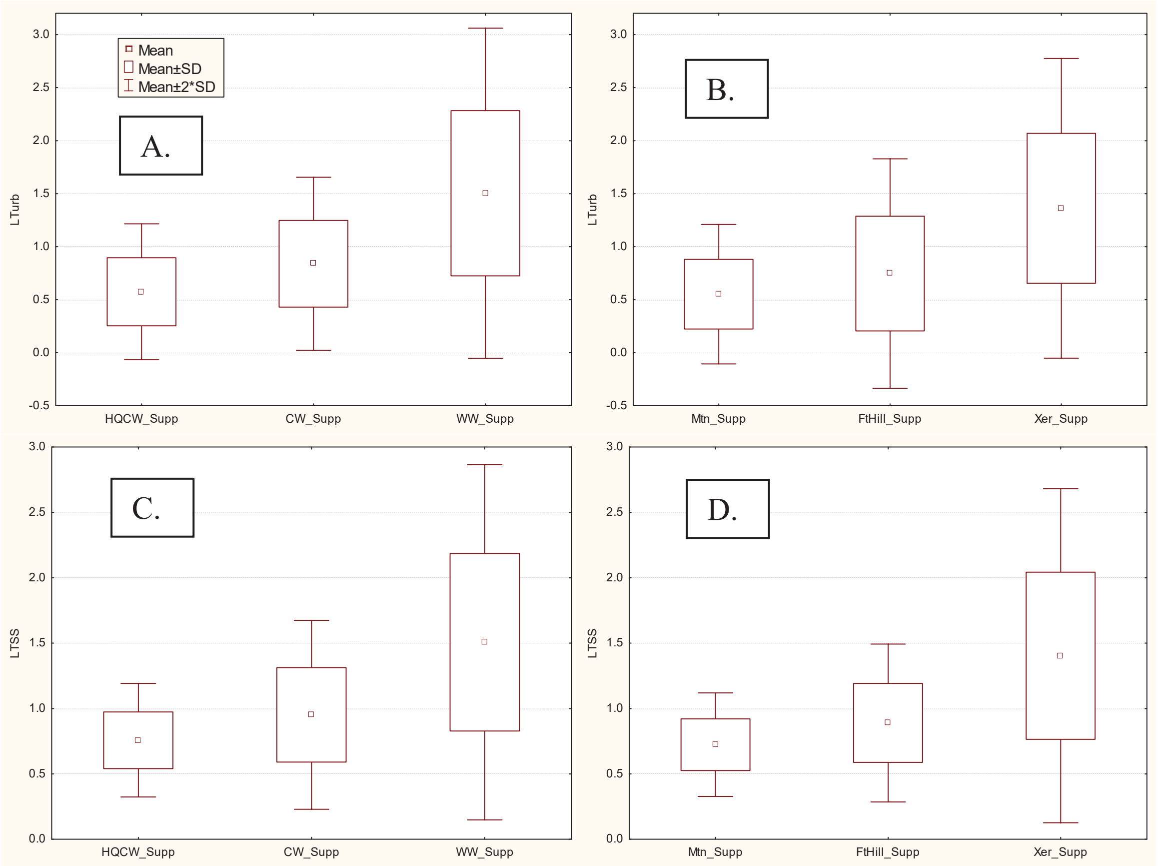

Figure 11. Suspended sediment value distributions for turbidity (A, B) and TSS (C, D) in sites supporting their uses in ALU classes (A, C) and bedded sediment classes (B, D).

Designated Uses

Aquatic life uses are the designated uses that this report addresses in establishing potential sediment benchmarks. Because the scientific literature is replete with examples of sediment effects on stream biota (Waters 1995, Wood and Armitage 1997, Suttle et al. 2004, Opperman et al. 2005), we assume that each of the indicators is associated with the designated use, satisfying the first selection criterion (a) listed above. Each of the indicators being considered is a direct measure of either absolute sediment conditions or conditions relative to expectations for the environmental setting. We confirmed this assumption of biological effects in our examination of sediment-response relationships (**Section 2.5**).

Data Availability and Analytical Appropriateness

We had sufficient data in and around New Mexico for analysis of all indicators listed in **Table 2** and their derivatives based on environmental settings. The fewest samples were from targeted riffle sites, which may limit the certainty of some of those analyses. However, the targeted riffle sites were important in determining the sensitivity of the methods, so we did not eliminate them. Particle embeddedness was originally considered as a potential indicator, but it was abandoned because the workgroup had a general lack of confidence in the consistency of the methods used in its measurement.

All of the indicators were appropriate for the analyses we considered. Where data gaps existed in ancillary data, those data were removed from analysis or, if appropriate, mean values were substituted for missing values. For analyses that required normal distributions (e.g., PCA, Pearson correlations), variables were transformed using the logarithmic or Arc Sine – Square Root functions. Sediment indicators were averaged over time at individual sites when data from multiple sampling events were available. For suspended variables, averaging only occurred when there was an indication that the flow conditions among samples were similar (low, runoff, or storm flows). Flow conditions were assigned using best professional judgment based on 1) date of sampling, 2) available flow data from the sampling event, and 3) field notes vs. an analysis of site hydrographs because the vast majority of sites with suspended solids data did not have associated stream gage data.

Measurement Reliability and Applicability

Some indicators require more effort in the field for accurate measurement. Percent sand & fines, % fines, and turbidity are among the simplest indicators to sample. The RBS and complex adjustments of the percentage measures (see Section 2.4.1) require more stream channel measurements or watershed characterization. TSS measurement is a somewhat more complex procedure than turbidity, requiring laboratory processing. In addition, turbidity is sensitive to organic materials and water coloration to a greater degree than TSS. For these reasons, much of the indicator characterization focused on comparisons among indicators. If simpler measures could substitute for more complex ones without sacrificing assessment accuracy, the simpler measure might be adequate and more efficient. If more complex measures were necessary for the most accurate assessments, then the simple measures may only suffice for screening applications.

2.4.1 Bedded Sediment Indicator Characteristics

RBS has proven to be a sensitive and meaningful indicator in other studies (Jessup 2009a, Kaufmann et al. 2009, Kaufmann and Hughes, 2006). Because fluvial site conditions are major determinants of the substrate conditions in stream channels, the critical particle size calculated from fluvial characteristics is a predictor of dominant and stable substrate conditions. The critical particle size is calculated from channel dimensions, roughness factors, and shear stresses (Kaufmann et al. 2008). The value is used as the denominator in the LRBS. Sediment conditions relative to the fluvial potential are better estimates of system stability and imbalance than absolute measures of fine sediment concentration because they intrinsically account for site-specific natural settings.

Preliminary analyses suggested that LRBS calculated without bedrock or hardpan components (termed LRBS_NOR) might improve associations between the bedded sediment measure and biological responses. The LRBS_NOR was a measure that regarded only the potentially mobile streambed particles in determining the geometric mean particle size. Calculated values related to the mobile particle distribution were missing about 25% of the records. In those cases, LRBS_NOR was estimated from LRBS (code = LRBSfin) and the percentages of bedrock and hardpan. Only 10% of cases required any adjustment because 15% had no bedrock or hardpan.

In contrast to LRBS, the percent fines and percent sand & fines measures are absolute quantities, which, except for natural variability captured by site classification, are more susceptible to natural variations. To decrease the variability and increase the specificity of percent fines and percent sand and fines responses to anthropogenic disturbances, they were adjusted to site-specific conditions in two ways. First, the observed values in reference sites of the active dataset were regressed on the log-critical particle size (code = ldcbf_fin), establishing the expected relationship in sites with minimal disturbance. Residuals from the reference regression were calculated to determine the degree of fine sediment accumulation relative to the expectation based on critical particle diameter. A constant was added to the residual to place most values in the positive scale (though this does not imply a relation to the absolute percentage vales). The critical particle size can only be calculated from reachwide data (not targeted riffle data). The terms used to describe these adjusted values were Resid_pSAFN and Resid_pFN.

$$\text{Resid\_pSAFN} = \text{pct\_SaFn} - (68.5599 - 21.8743 * \text{ldcbf\_fin}) + 20$$

$$\text{Resid\_pFN} = \text{pct\_Fn} - (48.1525 - 17.2526 * \text{ldcbf\_fin}) + 20$$

The second adjustment to the percentage indicators was based on data from 1,078 wadeable and non-wadeable streams and rivers in the western U.S. (Stoddard et al. 2005). Two multiple regression models were developed for % fines and % sand & fines. They were both based on bankfull bed shear stress, catchment mean annual precipitation, EPA aggregated ecoregion (Mountain, Xeric, or Plain), and the method used to quantify the bed substrate (wadeable stream method or non-wadeable river method). The residual of % sand & fines (Res2pSAFN) was calculated as Pct_SAFN – (Expected PCT_SAFN), where the Expected value of Pct_SAFN was calculated from the following regression model:

$$\begin{aligned}\text{Expected Pct\_SAFN} = {} & 56.13 - 13.15(\text{LDMB\_BW5}) - 35.07(\text{LPRECIPM}) - \\ & 8.30(\text{ECO3\_MT}) + 26.05(\text{ECO3\_PL}) - 12.66(\text{RIVER}).\end{aligned}$$

Where:

LDMB_bw5 = Log10 of bankfull bed shear stress,
LPRECIPM = Log10 of 30 year mean annual precipitation at centroid of catchment,
ECO3_MT = 1 if ECO3=MT, otherwise = 0,
ECO3_PL = 1 if ECO3=PL, otherwise = 0,
RIVER = 1 if non-wadeable river field methods were used, otherwise = 0.

The residual of % fines (Res2pFN) was calculated as Pct_FN – (Expected Pct_FN), where the Expected Pct_FN was calculated from the following regression model:

$$\begin{aligned}\text{Expected Pct\_FN} = {} & 35.6 - 9.55(\text{LDMB\_bw5}) - 23.9(\text{LPRECIPM}) - 5.22(\text{ECO3\_MT}) + \\ & 18.5(\text{ECO3\_PL}) - 13.2(\text{RIVER})\end{aligned}$$

Predictor variables were the same as for Expected Pct_SAFN. The model for % sand & fines had $R^2=0.54$ and RMSE=22.0, with $p<0.0001$ for the model and all predictors. The model for % fines had $R^2=0.37$ and RMSE=21.9, with $p<0.0001$ for the model and all predictors. Residual values near zero indicate conditions as would be expected with minimal disturbance and extremely positive values indicate excessive fine sediments.

In addition to the indicator adjustments described above, we continued to analyze the absolute percentages of fines and sand & fines, because regardless of the refined adjustments to specific fluvial site conditions, the biological assemblages may actually respond more clearly to the absolute values. Relative stability may not be relevant to biological impairment because biota adjust to existing substrate conditions, regardless of whether it is “supposed” to be there or not. In other words, do we expect the same biological community in two streams that have 30% sand & fines when one stream is stable and the other is unstable?

Targeted riffle samples are of interest because they coincide with the typical benthic samples collected by NMED from 1998 - 2005 (except for select sites from 1999 to 2001) because the filling in of interstitial spaces in biologically-important habitat components such as spawning gravels typically found in riffle areas is known to impact associate biota (Chapman and McLeod 1987, Lisle 1989, Waters 1995). This data collection technique did not include sampling outside of a representative riffle habitat. Targeted riffle data are somewhat easier (more efficient) to sample because they do not require transect sampling throughout the entire reach. The LRBS measure cannot be calculated from targeted riffle samples, so it is only the % fines and % sand & fines that are under consideration in this comparison. The questions posed by NMED and that we tried to answer through comparative analyses include:

> Are the sediment measures collected using targeted riffle sampling methods as sensitive as the sediment measures collected reachwide, or are they sensitive enough to detect sediment imbalances that would affect aquatic life? In other words, what is gained through reachwide sampling that is not possible with targeted riffle sampling (in the context of evaluating impairment)?

In order to generate a dataset to attempt to answer these questions, NMED performed pebble counts in both targeted riffle areas and reachwide during 2006 and 2007. We tried to answer these questions by analyzing the responsiveness of the targeted riffle samples to the general stressor gradient and the relationship of targeted riffle sediment indicators to biological conditions.

2.4.2 Bedded Sediment Reference Conditions

Percentiles of the reference distributions can be used as one line of evidence for establishing sediment benchmarks. Percentiles that are used to compare sites to reference conditions usually are from the median to the 90th percentile for values that increase with disturbance, such as % sand & fines. Below the median, conditions are similar to the best half of the reference sites. Values above the 90th percentile are unlike most of the reference sites and 10% of reference sites might have high values due to sampling and natural variability. Typically, quartiles are used to define similarity to or separation from the reference condition (Barbour et al. 1999, U.S. EPA 2000, U.S. EPA 2006).

Using the classification scheme to reduce natural variability of the indicators, the sediment conditions in reference sites were established. The reference sediment conditions are expected at all minimally disturbed sites of the same site class. Potential benchmarks (**Table 10**) were derived from the 10th and 25th percentiles for the LRBS indicator that decreases with increasing disturbance, and the 75th and 90th percentiles for the percentage indicators that increase with disturbance. Selection of benchmarks using these percentiles has precedence in state and federal biological and nutrient criteria programs (U.S. EPA 2006).

Percentiles of the indicators that can discriminate reference from stressed conditions may carry more weight in benchmark selection because the discrimination is evidence that the sediment measure actually varies with the landscape-level stressors that were used to define reference and stressed sites (see **Section 2.2**). Discrimination efficiency (DE) is a measure of the difference in values from reference sites to stressed sites. It is measured as the percentage of sites in stressed sites with indicator values worse than the 75th percentile of reference values. For the LRBS, stressed sites values below the 25th percentile of reference were used because these values are typically lower with greater stress. Box plots of the indicators in reference categories (**Appendix C**) were used to estimate the DEs listed in **Table 10**.

From the statistics in **Table 10**, we can see that the % sand & fines indicator was among the best at discriminating stress in each region. Percent fines performed similarly, except that it performed poorly in the Xeric region. Residuals of % sand & fines performed well in the Xeric regions. Residuals of % fines performed well in the Mountain and Foothill regions. LRBS was not best at discriminating stress as we have defined it here. Percent sand & fines performed well in targeted riffles wherever data were available to determine the DEs (for example, data were not available for the Xeric region). For the unadjusted percentage measures, the 75th percentile values of reference are within expected ranges in the Mountains and Foothills. They are relatively high in the Xeric region, but may not be unreasonable for that site class. LRBS 25th percentile of reference values are as expected and equal in the Mountains and Foothills.

Partial Correlations

To help in determining which of the bedded sediment indicators are most responsive to stressors, a partial correlation analysis was performed including all sites. The effects of catchments area, site elevation, stream slope, precipitation, and stream power were controlled so that the remaining variability would be more directly attributed to the human-influenced variables. The partial correlations that were significant ($p<0.05$) and relatively strong ($r^2 > 0.09$) were summarized by site class, indicator, and stressor (**Table 11**).

Table 10. Bedded sediment indicator statistics for reference sites in three site classes. Codes are as described in Table 6.

Indicator	Ref N	Mean	Min	10%ile	25%ile	Median	75%ile	90%tile	Max	Std.Dev	DE[b]
Mountains											
PCT_SAFN	55	16	0	2.2	5.6	13.3	20.6	35.1	65.7	14.1	>50%
PCT_FN	55	9.7	0	0	2.8	5.7	12.9	24.6	61.9	11.7	>50%
LRBS_fin	55	-0.7	-2.2	-1.5	-1.1	-0.7	-0.4	-0.1	0.5	0.6	25-50%
LRBS_NOR	55	-0.8	-2.2	-1.5	-1.1	-0.8	-0.4	-0.1	0.4	0.5	25-50%
trPCT_SAFN[a]	5	7.8	3	3	3	7	12	13.2	14	5.1	>50%
trPCT_FN[a]	5	0.8	0	0	0	1	1	1.6	2	0.8	50%
Resid_pSAFN	55	14.6	-8	1.6	6.4	13.6	22.8	27	58.5	12.2	50%
Resid_pFN	55	18.8	-4.8	8.1	11.8	18.6	23.5	28.9	49.5	9.4	>50%
Res2pSAFN	55	-21.8	-38.1	-34.7	-29.2	-24	-17.8	-9.9	17.4	11.7	50%
Res2pFines	55	-12.6	-23.6	-22.1	-18.4	-13.6	-9.8	-2.1	31.9	9.7	50%
FootHills											
PCT_SAFN	27	27.7	0	5.1	18.6	27.1	36.9	45.2	61	15.5	<75%
PCT_FN	27	11.2	0	0.9	2.4	10.9	18.6	23.7	31.8	9.6	75%
LRBS_fin	27	-0.8	-2	-1.6	-1.1	-0.7	-0.4	0	0.3	0.6	>50%
LRBS_NOR	27	-0.9	-2.0	-1.7	-1.3	-1.0	-0.5	-0.1	0.2	0.6	50%
trPCT_SAFN[a]	8	9.5	0	2.8	5.1	8	11.4	20.6	22	7.6	>75%
trPCT_FN[a]	8	1.4	0	0.4	0.9	1	2.3	3	3	1.1	50%
Resid_pSAFN	27	19.3	-8.6	2	7.4	17.5	30.3	41.1	51.7	16	>50%
Resid_pFN	27	14.8	-4.4	3.8	10	13.1	18.5	28.2	43.5	10.1	<75%
Res2pSAFN	27	-21.2	-46.9	-38.4	-34.8	-20.8	-11.3	-0.8	8.9	14.6	>50%
Res2pFines	27	-18.8	-36.1	-28.7	-25.2	-20.9	-12.8	-3.9	5.1	9.7	75%
Xeric											
PCT_SAFN	17	57.2	19.5	29.8	43.8	53.3	74.3	84.4	99	22.8	>50%
PCT_FN	17	38.7	4.8	8.2	18.1	34.3	59	72.2	86.7	26	<25%
LRBS_fin	15	-1.8	-2.8	-2.5	-2.3	-1.9	-1.3	-0.9	-0.2	0.7	25-50%
LRBS_NOR	15	-2.0	-2.9	-2.7	-2.5	-1.9	-1.6	-1.1	-0.7	0.7	25-50%
trPCT_SAFN[a]	6	17.7	0	2.5	6.8	13	27.5	37.5	43	16.5	NA
trPCT_FN[a]	6	7.3	0	0	0	5.5	13.3	16.5	19	8.4	NA
Resid_pSAFN	15	43.9	10.9	18	27.9	41.2	57.5	64.4	82.4	20.7	>50%
Resid_pFN	15	36.5	-0.6	11.7	16.2	32.3	57.2	68.8	71	23.4	<25%
Res2pSAFN	15	-18.1	-57	-51.9	-31.6	-17.7	-3.1	12.7	26.3	24.7	50%
Res2pFines	15	-11.2	-42.2	-39.9	-25.6	-11.8	2.6	19.4	20.8	21.5	<25%

NOTES: NA = Insufficient data from stressed sites.

[a] Targeted riffle data. The number of reference sites (“ref N”) for trPCT_SAFN and trPCT_FN in the three site classes is low compared to the other indicators because these data are not part of the standard EMAP data collection and were only collected at EMAP sites sampled by NMED from 2006-2007.

[b] DE = Discrimination Efficiency

Table 11. Bedded sediment indicator / stressor relationships that were significant and relatively strong in partial correlation analysis, all sites.

Indicator	Class	Variable	Partial r	Indicator
% fines	ArcSin(%Fn)	Mtn	COND	0.408***
		FtHill	COND	0.326*
		FtHill	Dams_km2	0.346**
		FtHill	W1_HAG	0.327*
		Xer	Dams_km2	-0.409**
	Res2pFines	Mtn	COND	0.463***
		FtHill	Dams_km2	0.343**
		FtHill	W1_HAG	0.325*
		Xer	COND	0.329*
	Resid_pFN	Mtn	COND	0.465***
		FtHill	Dams_km2	0.362**
		FtHill	W1_HAG	0.378**
		Xer	Dams_km2	-0.320*
	ArcSin(%Fn) in targeted riffles	ALL	%HiErod	0.326*
		ALL	COND	0.787***
		ALL	DO_FLD	0.351*
		ALL	LRdX_km2	0.364*
		ALL	PFORnew	-0.440**
		ALL	W1H_CROP	0.482**
		Xer	COND	0.933***
LRBS	LRBS_fin	Mtn	COND	-0.481***
		FtHill	Dams_km2	-0.321*
		FtHill	W1_HAG	-0.421***
		Xer	COND	-0.364**
	LRBS_NOR	ALL	COND	-0.322***
		Mtn	COND	-0.500***
		FtHill	%ModHiErod	-0.301*
		FtHill	Dams_km2	-0.359**
		FtHill	W1_HAG	-0.409**
		Xer	COND	-0.319*
% sand & fines	ArcSin(%SaFn)	ALL	COND	0.363***
		Mtn	COND	0.606***
		FtHill	%ModHiErod	0.370**
		FtHill	COND	0.323*
		FtHill	W1_HAG	0.311*
		Xer	COND	0.339*
	Res2pSAFN	Mtn	COND	0.629***
		FtHill	%ModHiErod	0.344**
		FtHill	W1_HAG	0.318*
		Xer	COND	0.360**
	Resid_pSAFN	ALL	COND	0.331***
		Mtn	COND	0.588***
		FtHill	%ModHiErod	0.302*
		FtHill	W1_HAG	0.384**
		Xer	COND	0.306*
	ArcSin(%SaFn) in targeted riffles	ALL	COND	0.734***
		ALL	DO_FLD	0.349*
		ALL	LRdX_km2	0.3689*
		ALL	PFORnew	-0.425**
		ALL	W1H_CROP	0.432**
		FtHill	W1_HAG	0.726*
		Xer	COND	0.887***

The partial correlations that were significant in all sites, but not in the individual site classes suggest that the stressors intensities are accounted for by site classes. The targeted riffle sample sizes were small in the individual site classes, which could lead to spurious results. The consistent relationships appear to be with the stressors conductivity, dam density, and riparian land use. Percent sand & fines were also related to the percent moderately and highly erodible lithology in the Foothills site class. All versions of the percentage indicators had similar responses.

2.4.3 Comparisons for suspended measures

Turbidity and TSS values were compared in sites where both were measured. Regression analysis and ANOVA were performed within three ALU categories and on a log-log scale. The regression analysis yielded the slopes and intersects, with slopes near 1.0 and intersects near 0 indicating similar measurement values. The ANOVA was used to find the root-mean-squared-error (RMSE), which is an estimate of the error in translating one value to another. The RMSE can be applied to individual observations to estimate the 1 standard-deviation error bounds. Application begins with the antilog of the RMSE, which is divided into the observed value to find the lower bound and multiplied by the observation for the upper bound.

We limited the comparison to low flow conditions because these were the most commonly collected data, low flow conditions are the assumed applicable benchmark conditions for this analysis, and high outliers that might skew comparisons can be limited. Samples were designated as low flow samples were confirmed by comparing average TSS and turbidity values across sites. Data were plotted in bi-plots of turbidity versus TSS and in box plots. Outliers on both of these plots were scrutinized and either eliminated or corrected. Corrections usually involved re-designation of samples as storm-flow samples when actual flow data were missing.

Turbidity and TSS were correlated and had regression slopes near 1.0 in each of the ALU categories (**Figures 12-14**). Because the slopes were near 1.0 in each case, we can generalize that 1.0 ntu turbidity unit is similar to 1.0 mg/L TSS unit for this data set. In the HQCW category, the slope was 1.03 and the intersect was -0.14, indicating that for the low flow condition, TSS values were higher than turbidity value by 0.14 log units (1.38 TSS/turbidity units) over the entire range. In CW sites, the intersect was close to 0 and the slope was 0.94, showing that the bias towards higher TSS values is more prevalent when both values are higher. The regressed relationship in WW sites was closest to the 1:1 line, among ALU categories.

The regression coefficients, RMSE, and 1 standard deviation intervals within ALU categories and over all sites (**Table 12**) suggest that there is a potential for translating TSS values to turbidity values. In the HQCW category, the apparent bias should be taken into account; it is likely that the TSS value would be higher than the turbidity value. For all data combined, the 1 standard deviation confidence bound around 10 ntu is 5.89 – 17.0 ntu, which should be taken into account when making translations among ntu and mg/L units.

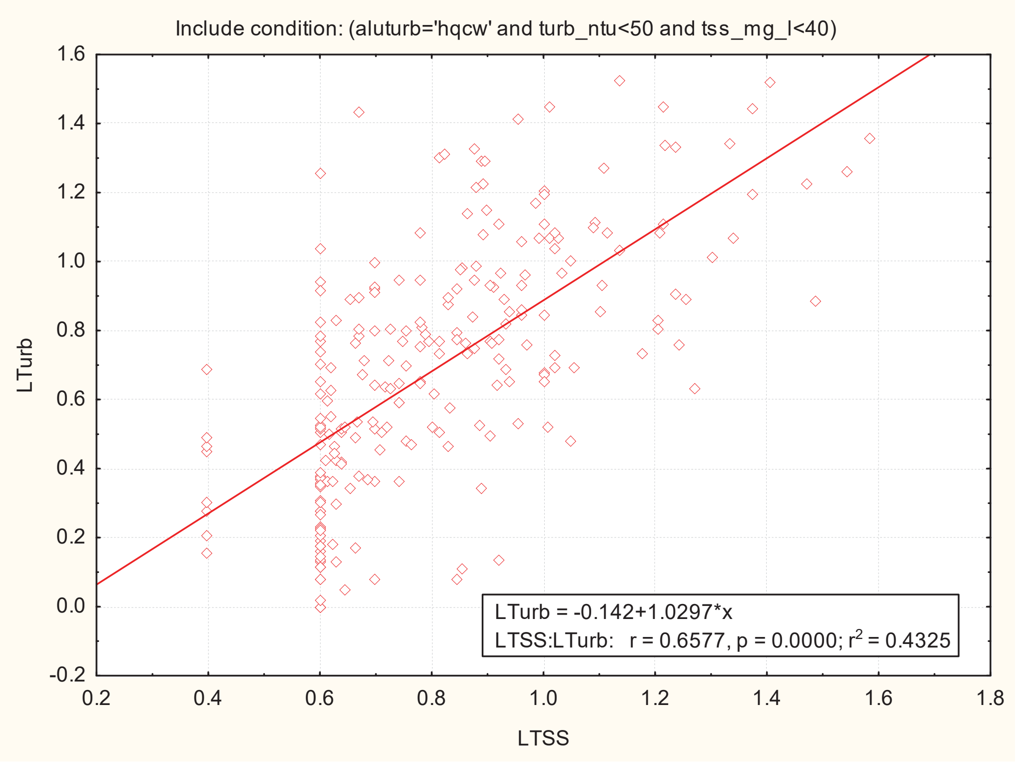

Figure 12. Relationship of turbidity to TSS (low-flow, log-transformed) in HQCW sites.

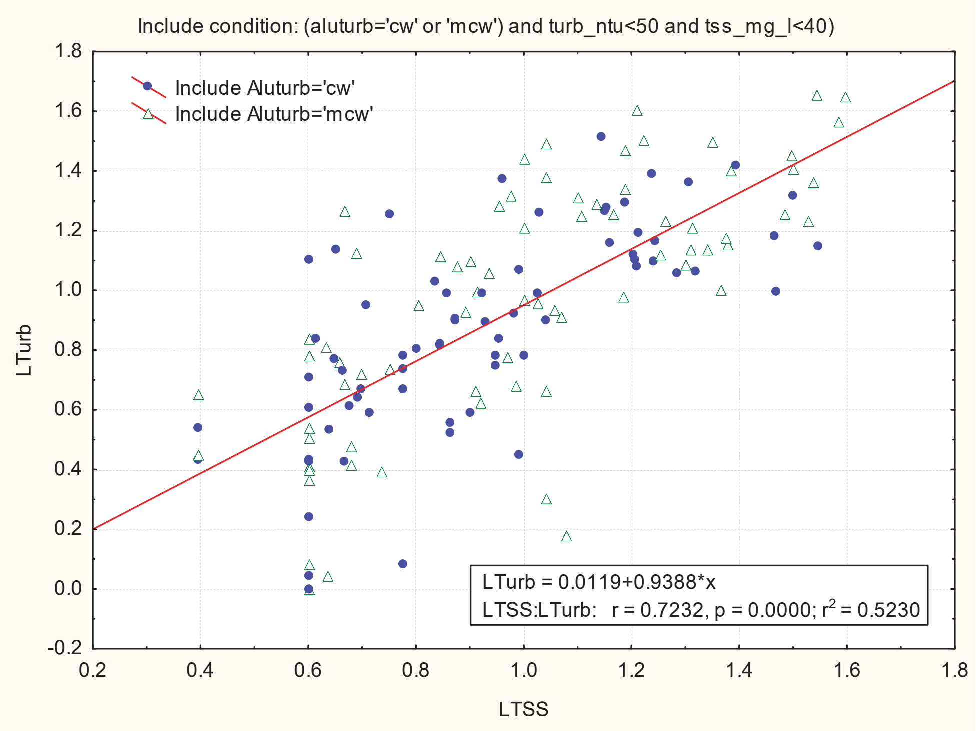

Figure 13. Relationship of turbidity to TSS (low-flow, log-transformed) in Cold-water (cw, blue dots) and Marginal Cold-water (mcw, green triangles) sites.

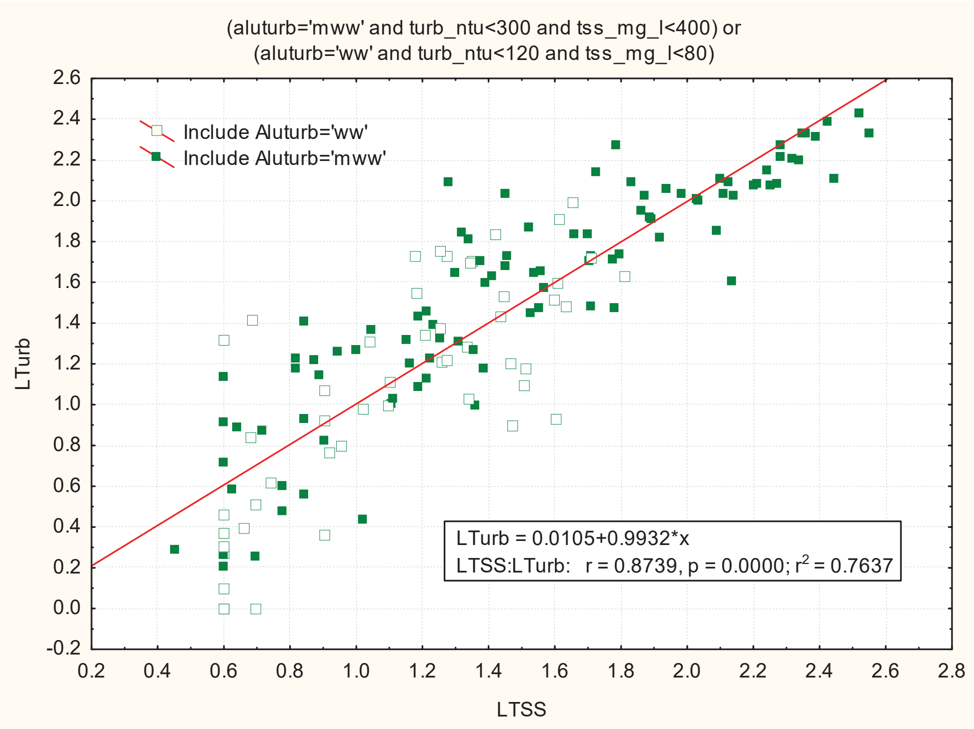

Figure 14. Relationship of turbidity to TSS (low-flow, log-transformed) in Warm-water (ww, empty squares) and Marginal Warm-water (mww, filled squares) sites.

Table 12. Comparative statistics relating turbidity and TSS.

ALU category	Regression equation for turbidity (log-log scale)	Regression coefficient (r^2)	RMSE	1 s.d. interval around 10 ntu
HQCW	-0.142 + 1.0297*TSS	0.4325	0.21	6.12 – 16.3
CW	0.0119 + 0.9388*TSS	0.5230	0.21	6.22 – 16.1
WW	0.0105 + 0.9932*TSS	0.7637	0.22	6.05 – 16.5
All sites	-0.0897 + 0.995*TSS	0.7507	0.23	5.89 – 17.0
All sites (values untransformed)	5.1755 + 0.7814*TSS	0.7379	41	-31 - 51

Comparisons of suspended measures to bedded measures

Suspended sediment measures can be interpreted in the context of both flow conditions and substrate conditions. In this analysis, we made a simple distinction between mountainous and non-mountainous regions, where mountains include the Southern Rockies, the Arizona/New Mexico Mountains, and Madrean Archipelago (ecoregions 21, 23 and 79). We present turbidity and % sand & fines because they are the indicators used in current WQS. Correlation analysis revealed that in low-flow conditions, turbidity was more highly correlated with % sand/fines than with LRBS or % fines. In bi-plots, turbidity appears more strongly correlated to bedded measures than does TSS. Additional plots illustrating additional sediment indicators and the correlation table are included in **Appendix D**.

By plotting turbidity in relation to % sand & fines for both low and storm flows it is clear that with less fine sediment in the streambed there is less likelihood of elevated suspended sediments (**Figure 15**). When % sand & fines in the streambed were < 20% (or LRBS > -0.5), turbidity was consistently less than 10 ntu. During storm flows, suspended measures were consistently higher than hypothetical benchmarks for low flows. However, the higher suspended sediment values were observed in sites with fine and unstable sediments. These relationships may be instructive when considering different types of stresses associated with bedded and suspended sediments.

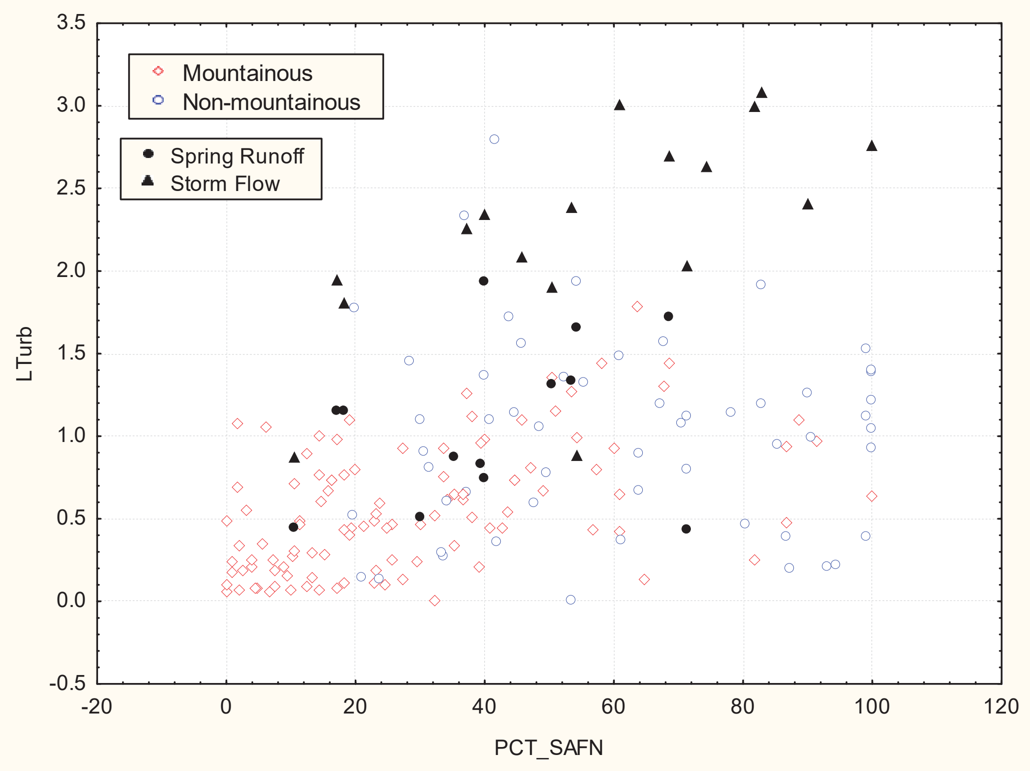

Figure 15. Suspended sediment as a function of % sand & fines.

2.4.4 Suspended Sediment Percentiles

As described in the **Section 2.2**, we did not describe true reference conditions for the suspended sediment data set. Instead, in most cases we used the assessment of ALU support as a measure of relative disturbance. These assessments of ALU support may or may not have included sediment information, which may bias statistics of sediment conditions in Fully Supporting sites because the sediment stress levels had been pre-determined. Therefore, we present the indicator percentile values in all sites as well as in Fully Supporting sites. The percentile values from all sites have been used in similar criteria identification approaches (e.g., EPA nutrients) and are not biased by potentially improper identification of the disturbance gradient.

The greatest differences in the distributions are among site classes, with less difference between ALU support statuses within the individual classes (**Figure 16**). The statistics shown (**Table 13**) may inform selection of sediment benchmarks. Because 1) the discrimination of supporting and non-supporting sites is only marginally effective with these indicators and 2) paired flow data at the time of sampling along with site hydrograph data to were not available, selection of any percentile as a benchmark should be aided by stressor-response relationships, fish and other biota tolerance values in available literature, and professional judgment.

2.5 Biological Responses to Sediment Conditions

Quantile regression is a method for estimating functional relations between variables along the upper boundary of the conditional distribution of responses (Cade et al. 1999). If limiting factors such as sediments act as constraints on organisms, then the estimated effects for the measured factors are related to some upper limit. This is apparent when the biological measures tend to exhibit an upper limit that varies with the value of a disturbance variable — that is, the maximum biological condition generally falls beneath a sloping line in a scatter plot of biological condition against the disturbance variable. Points that are not along the slope (in the heel of the wedge) represent sites with worse biological conditions due to factors not represented on the *x*-axis. The slope represents biological potential (plotted on the *y*-axis) in relation to the disturbance of interest (plotted on the *x*-axis). Estimation of the limiting slope is accomplished through quantile regression, which was performed using R software (R Development Core Team 2010) and associated code (quantreg).

The quantile regression analysis was conducted such that several upper quantiles (75th, 85th, 90th, and 95th) were calculated and plotted. When the upper quantiles are relatively parallel, the biological potential is likely limited by the stressor variable (Cade et al. 1999, Bryce et al. 2008). This tells us there is a likely effect, but the point at which the effect becomes critical cannot be directly determined. The multiple upper quantiles were examined and parallelism was determined based on professional judgment regarding the consistency of the slopes and the meaningfulness of the 90th quantile regression line (good, flat, or inconsistent). When the 90th quantile regression line was good, it was plotted to illustrate the change in a biological resource for each increment of sediment disturbance.

The change-point is the point along an environmental gradient at which there is a high degree of change in the response variable. The data are divided into two groups, above and below a potential sediment benchmark, where each group is internally similar and the difference among groups is high. To determine the change-point, we use nonparametric deviance reduction (Qian et al. 2003, King and Richardson 2003) to identify thresholds in biological responses to sediments. This technique is similar to regression tree models, which are used to generate predictive models of response variables for one or more predictors. Using this comparison, the change-point is the first split of a tree model with a single predictor variable (i.e., fine sediment percentage). Change-points and statistical significance were obtained using R software (R Development Core Team 2010) and associated code (chngp.nonpar).

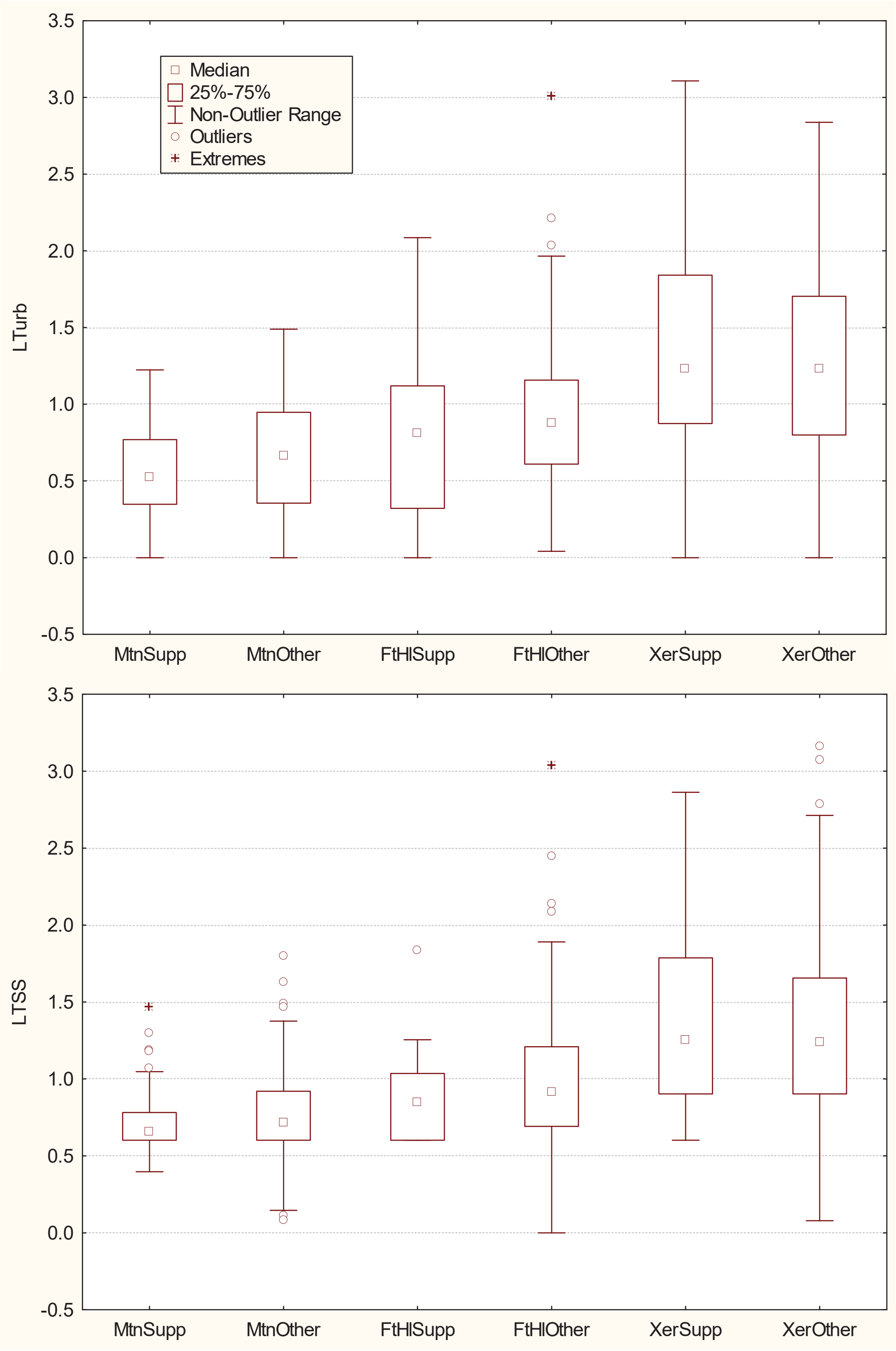

Figure 16. Suspended sediment distributions in fully supporting sites and other (not fully supporting) sites in three site classes.

Table 13. Suspended sediment indicator percentiles for fully supporting sites and all sites in three site classes.

		Fully Supporting Sites			All Sites		
		Valid N	75th	90th	Valid N	25th	Median
Mountains	Turbidity (ntu)	68	4.88	9.50	217	1.25	3.10
	TSS (mg/L)	70	5.05	8.75	221	3.00	3.89
FootHills	Turbidity (ntu)	24	12.18	19.30	136	2.33	5.99
	TSS (mg/L)	24	9.88	16.12	138	3.71	6.71
Xeric	Turbidity (ntu)	83	68.50	191.76	289	5.60	16.00
	TSS (mg/L)	85	60.23	262.80	295	7.00	17.00

One caveat of the change-point analysis is that a change-point may be identified, and even determined to be statistically significant, when the change-point value is actually only an artifact of the analysis and not an indication of a change in system properties. The method always finds a change-point, even in a dataset with a perfect straight line relationship between *X* and *Y*. It has been well established that sediment size affects macroinvertebrate assemblage characteristics (e.g., Waters 1995, Wood and Armitage 1997, Suttle et al. 2004, Opperman et al. 2005). Therefore, it is reasonable to believe an ecological threshold does exist between certain biological metrics and sediment conditions. In our analyses, we evaluated this relationship by examining the locally-weighted regression line (LOWESS or loess) fit on biplots of biological metrics and sediment indicators. If the LOWESS fit did not show a change-point, then the value identified through change-point analysis was disregarded.

The LOWESS technique (Cleveland 1979) is designed to address nonlinear relationships, which may be important when investigating changing responses along a stressor gradient. LOWESS combines the simplicity of linear least squares regression with the flexibility of nonlinear regression. It achieves this by fitting simple models to localized subsets of the data to build up a function that describes the deterministic part of the variation in the data, point by point. LOWESS fits segments of the data to the model, essentially, at the central tendency of the data. This method does not require specification of a global function of any form to fit a model to the data but to simply fit segments of the data to the model. We used a bandwidth that considered 75% of the data for smoothing the slope at each data point. The LOWESS regression line can be used in combination with other indicators of sediment effects, primarily as a visual confirmation of changing biological measures at certain sediment indicator values.

The relationships between bedded sediment indicators and biological metrics were examined using bi-plots showing significant change-points, meaningful quantile regression lines, the LOWESS regression line, and reference points in comparison to non-reference points. From these plots, we can discern three things:

1. When the sediment variable appears to be limiting the metric potential,
2. If the calculated change-point coincides with changes along the LOWESS regression line, and
3. If the change-point reasonably separates reference and non-reference points.

In **Figures 17 – 19**, significant change-points are shown as vertical lines, meaningful 90th quantile regression lines are shown as diagonal dash-dot lines, LOWESS regression lines are shown as non-linear solid lines, reference points are shown as solid circles, and non reference points are shown as crosses. In **Figure 17**, the LOWESS regression line is variable but high until about 15% sand & fines, where it begins to drop. The significant change-point at 20% sand & fines is at the part of the LOWESS line that drops below a previous trough. The two indications are in general agreement. In addition, the quantile regression lines are relatively parallel, strengthening our case that the stressor is limiting the biological potential. Reference points are more common to the left of the change-point. In all, the 20% sand & fines change-point is a potential benchmark.

Final recommendation of a biological threshold for an indicator in a site class would be based on corroborated results from all the metrics and the NMMSCI. In some cases, the lowest effect level may be appropriate to recommend as a benchmark. In other cases, the potential benchmarks indicated by significant change-points may be inappropriate (**Figure 18**) and benchmarks may be recommended based on other indications. Metrics that do not show a significant change-point or do not appear to be limited by the stressor (**Figure 19**) would not be used in determining a benchmark.

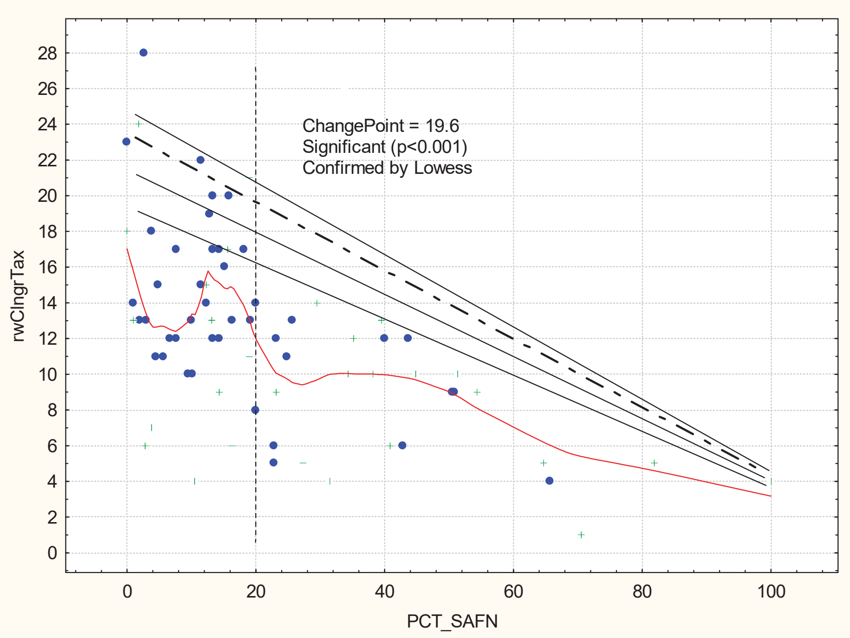

Figure 17. Example of significant change-point that is recognizable as a change in the y-axis value with respect to the x-axis. The LOWESS regression line shows an apparent shift near the change-point. The upper quantile regression lines are generally parallel, suggesting a true biological limitation. rwClngrTax = reachwide number of clinger taxa.

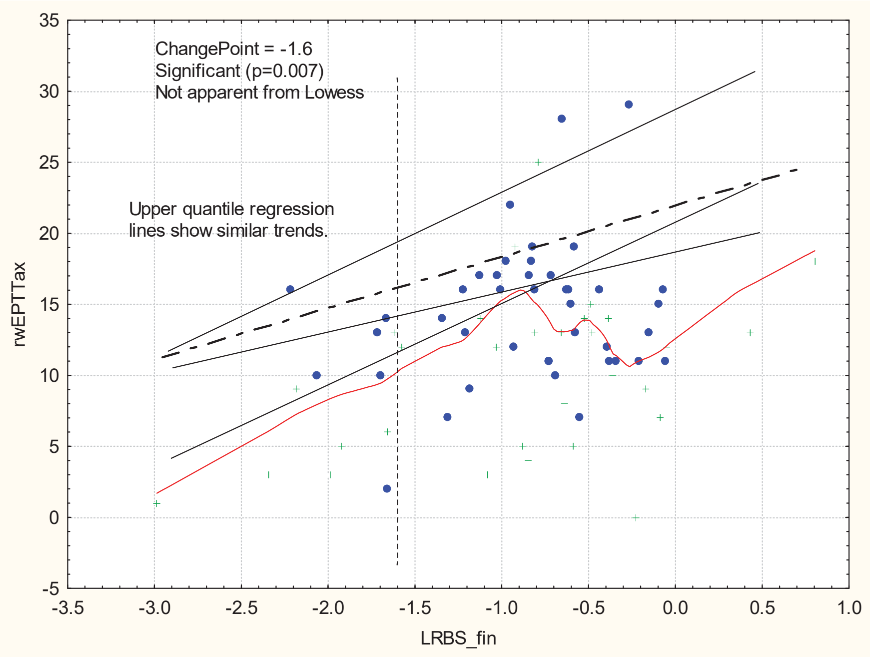

Figure 18. Example of significant change-point that does not "look" like an abrupt change-point. As suggested by the LOWESS regression line, it looks like a gradual decrease in metric values. The upper quantile regression lines are generally parallel, suggesting a true biological limitation. rwEPTTax = reachwide number of Ephemeroptera, Plecoptera, and Trichoptera taxa.

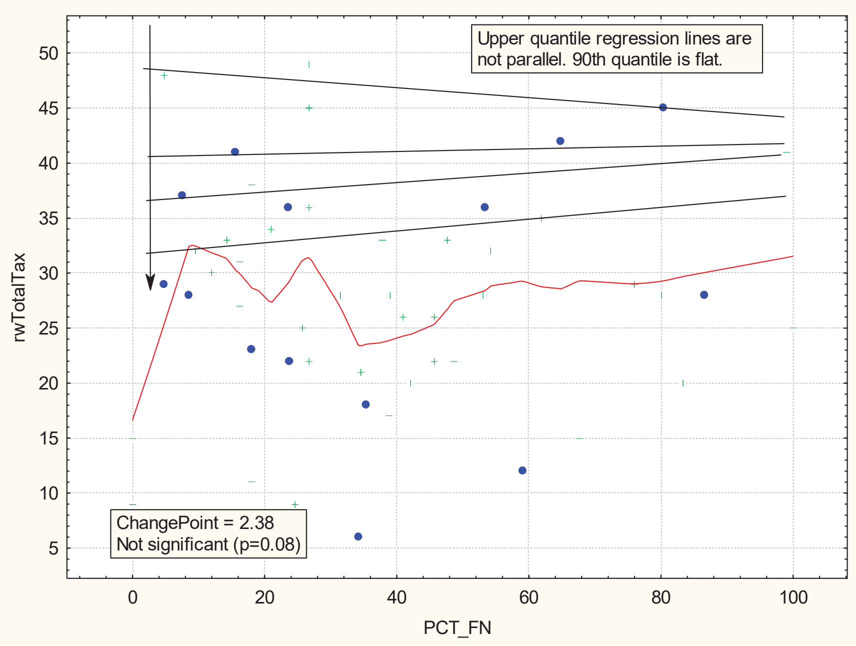

Figure 19. Example of a sediment-biological relationship that is indistinct. The change-point is not significant and the upper quantile regression lines have both positive and negative slopes.

2.5.1 Bedded Sediment and biological responses

The following plots are arranged by sites class (Mountains, Foothills, Xeric), metrics, and sediment indicators, which are as follows:

Metrics:

NMMSCI	New Mexico Macroinvertebrate Stream Condition Index
rw NMMSCIprop	NMMSCI as a proportion of the impairment threshold
rwTotalTax	Reachwide total number of taxa
rwEPTTax	Reachwide number of Ephemeroptera, Plecoptera, and Trichoptera taxa
rwEphemTax	Reachwide number of Ephemeroptera taxa
rw%sensEPT	Reachwide percent of sensitive EPT individuals
rwClngrTax	Reachwide number of clinger taxa
rwHBI	Reachwide Hilsenhoff Biotic index
rwIntolTax	Reachwide number of pollution intolerant taxa

Sediment Indicators:

PCT_SAFN	Percent Sand & Fines
PCT_FN	Percent Fines
LRBS_fin	Log Relative Bed Stability
LRBS_NOR	Log Relative Bed Stability after excluding bedrock and hardpan
Res2pSAFN	Residual of Percent Sand & Fines (Stoddard et al. 2005)
Res2pFN	Residual of Percent Fines (Stoddard et al. 2005)

In **Figures 20-25, 27-32, and 34-39**, significant change-points are shown as vertical lines, meaningful 90th quantile regression lines are shown as diagonal dash-dot lines, LOWESS regression lines are shown as non-linear solid lines, reference points are shown as solid circles, and non reference points are shown as crosses.

Mountains

Percent Sand & Fines (**Figure 20**): All of the metrics had meaningful upper quantile regressions, indicating that the biological condition is limited by increasing percentages of sand & fine sediments. In five of the eight metrics, the change-point was identified near 20%. Those metrics with change-points at higher levels also showed some variation in the LOWESS regression line at about 20%. While there were several reference sites with more than 20% sand & fines, most of them had less. Based on the 90th quantile regression line, an increase in 20% sand & fines results in a loss of five taxa.

Percent Fines (**Figure 21**): Seven of the eight metrics had meaningful upper quantile regressions, indicating that the biological condition is limited by increasing percentages of fine sediments. In all eight metrics, the change-point was identified near 20%. This is a little more than is indicated by the LOWESS regression shifts, which start at about 15%. Those metrics with change-points at higher levels also showed some variation in the LOWESS regression line at about 20%. While there were several reference sites with more than 20% fines, most of them had less than 15%. Based on the 90th quantile regression line, an increase in 20% fines results in a loss of four EPT taxa.

LRBS (**Figure 22**): Only three of the eight metrics had meaningful upper quantile regressions, probably because a high number of sites at the high end of the LRBS scale (>-1.0) show similar potential, or perhaps decreasing potential with the highest LRBS values. This is not unexpected because very high LRBS values indicate channel armoring, which can affect biota as much as channel instability. In four metrics, the change-point was identified near -1.6 units. The highest change-point was identified with the clinger taxa metric, which was near -1.1 units. The beginnings of the declines of the LOWESS regression lines are variable, between -1.3 and -0.9. While there were several reference sites with LRBS values below -1.6, most of them had values greater than -1.4 units. Based on the 90th quantile regression line, a decrease in 0.5 LRBS units results in a loss of two EPT taxa.

LRBS_NOR (**Figure 23**): After removing the immobile substrates from the bed stability equation, the LRBS NOR shows response patterns that are similar to the LRBS fin. Three of the eight metrics had meaningful upper quantile regressions and in four metrics, the change-point was identified near -1.6 units. The highest change-point was identified with the clinger taxa metric, which was near -1.0 units, which coincided with the beginnings of the declines of the LOWESS regression lines. Based on the 90th quantile regression line, a decrease in 0.5 LRBS units results in a loss of two EPT taxa.

Residual Percent Sand & Fines (**Figure 24**): Only two of the metrics had meaningful upper quantile regressions, indicating that the biological condition is limited for Ephemeroptera taxa and clinger taxa when sand & fines are high in relation to site-specific expectations. In four of the eight metrics, the change-point was identified near 19 units. The lowest change-point was identified with the clinger taxa metric, which was near -19 units. The LOWESS regression lines showed shifting trends at levels ranging from -20 to 0, suggesting that the median of significant change-points is too high and may be an artifact of the data distribution, in which the few sites with very high residual values have consistently very low metric values and at intermediate level, there are a few high outliers of the metrics. Most of the reference sites had residual values less than -4. Based on the 90th quantile regression line, an increase in 20 residual sand & fine units results in a loss of three EPT taxa.

Residual Percent Fines (**Figure 25**): All the metrics had meaningful upper quantile regressions, indicating that the biological condition is limited by increasing percentages of sand & fine sediments in relation to site-specific expectations. The change-point was identified in the range of -11 to -4 units, being lowest for the % sensitive EPTs and the HBI. The change-points generally coincided with the LOWESS regression lines, which showed shifting trends at about -10. Most of the reference sites had residual values less than 0. Based on the 90th quantile regression line, an increase in 20 residual fine units results in a loss of two taxa.

<u>Recommendations:</u>

Based on these analyses of stressor-response in the Mountains site class, the following values and ranges of values are recommended as potential benchmarks for preventing benthic macroinvertebrate degradation.

Percent sand & fines: At greater than 20% sand & fines, biological potential is limited.

Percent fines: While the change-points indicate a change at 20% fines, this is equal to the percentage for % sand & fines. Based on the LOWESS regression shift points and the fact that fines are a fraction of sand & fines, the recommended benchmark is 15% fines.

Log Relative Bed Stability: Based on effects to clinger taxa and the LRBS values in reference sites, the benchmark should be greater than is indicated by the median change-point (-1.6 units). The benchmark should be near -1.25 units (range -1.4 to -1.1), which coincides with the LOWESS declining trends and the reference site distribution

Log Relative Bed Stability NOR: Change-points identified for EPT taxa, clinger taxa, and the HBI are higher than the median -1.6 units, perhaps due to the more sensitive organisms represented in those metrics. The benchmark should be near -1.1 units (range -1.25 to -1.0), which coincides with the LOWESS declining trends and the reference site distribution

Residual of % sand & fines: Based on the effects on clinger taxa and LOWESS regression trends, the recommended benchmark for Residual of % sand & fines is -19 units. The regression of % sand & fines with the residual values shows that residual sand & fines of -19 units agrees with 20% sand & fines (**Figure 25**).

Residual of % fines: Based on the effects on sensitive EPTs, the HBI, and LOWESS regression trends, the recommended benchmark for Residual of % fines is -8 units. The regression of % fines with the residual values shows that residual fines of -8 units agrees with 15% fines (**Figure 25**).

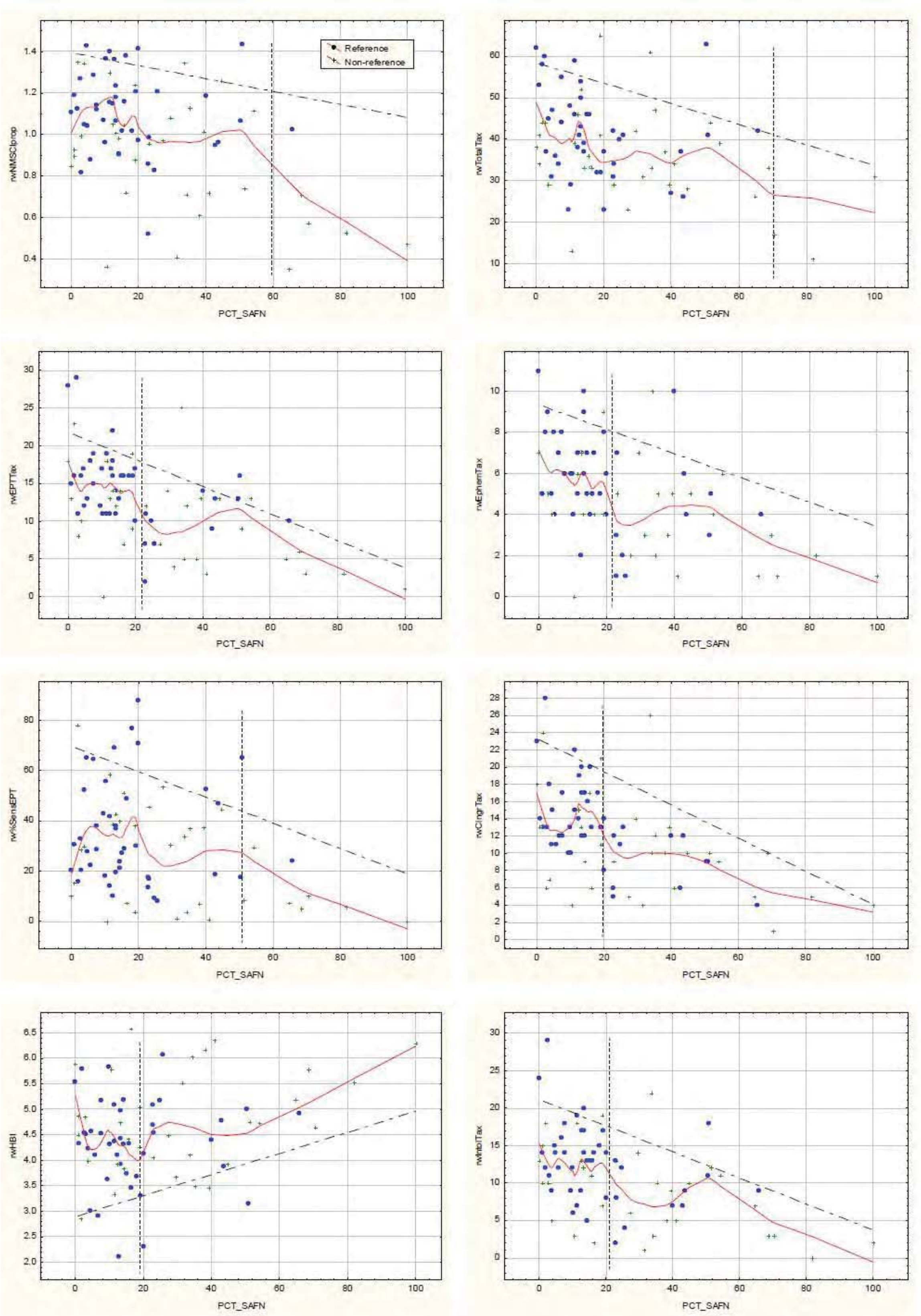

Figure 20. Biological responses to % sand & fines in the Mountain site class.

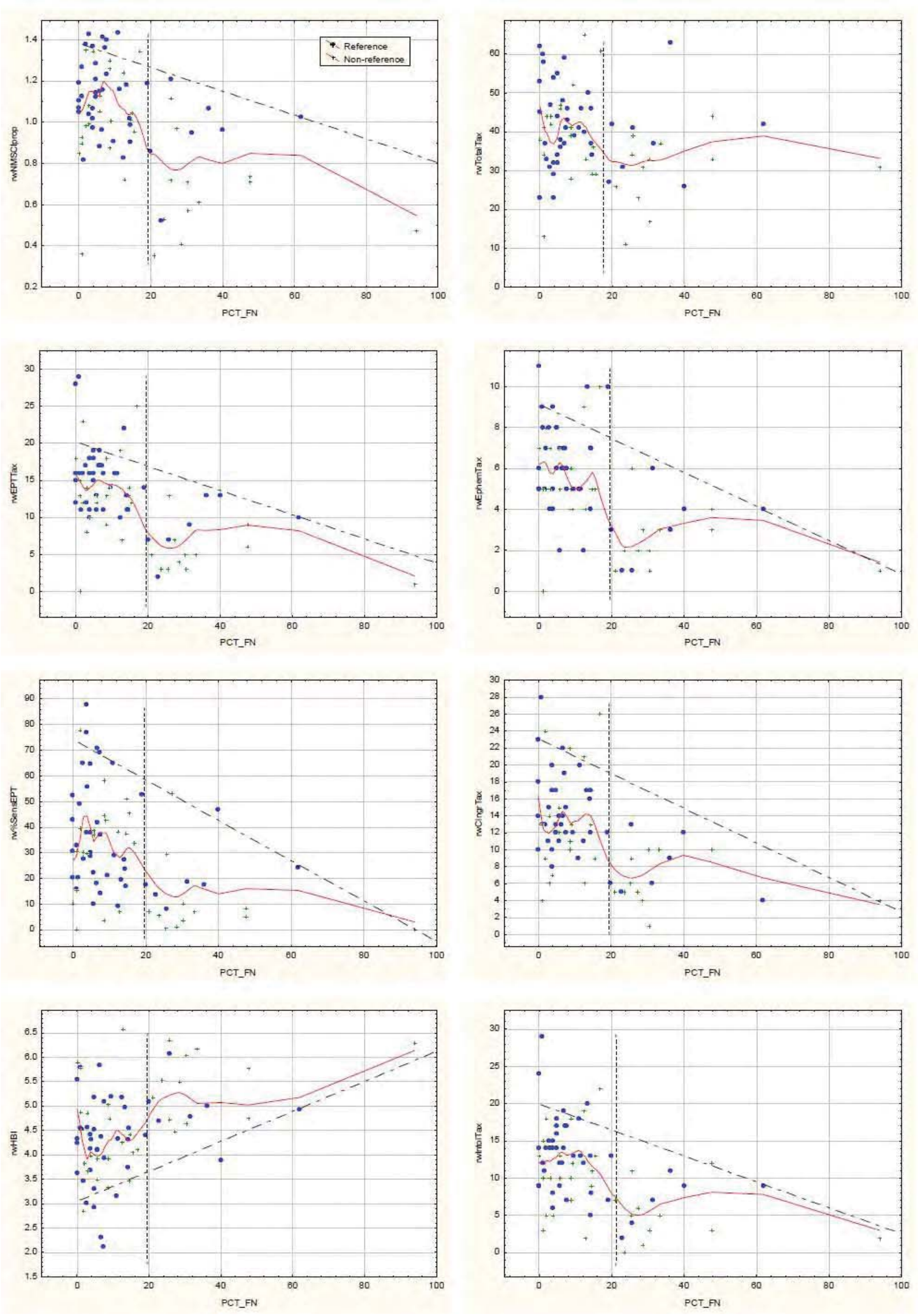

Figure 21. Biological responses to % fines in the Mountain site class.

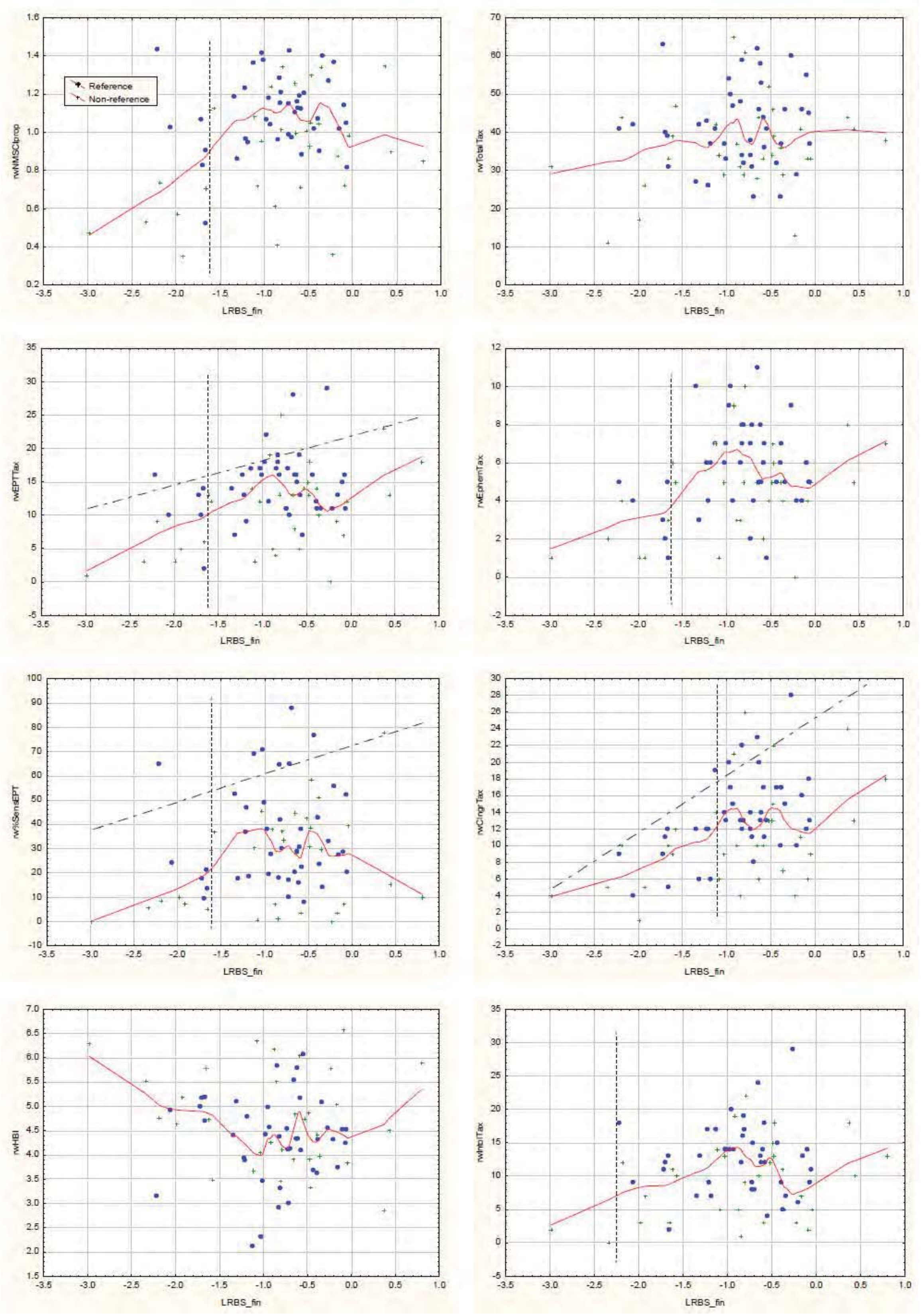

Figure 22. Biological responses to LRBS in the Mountain site class.

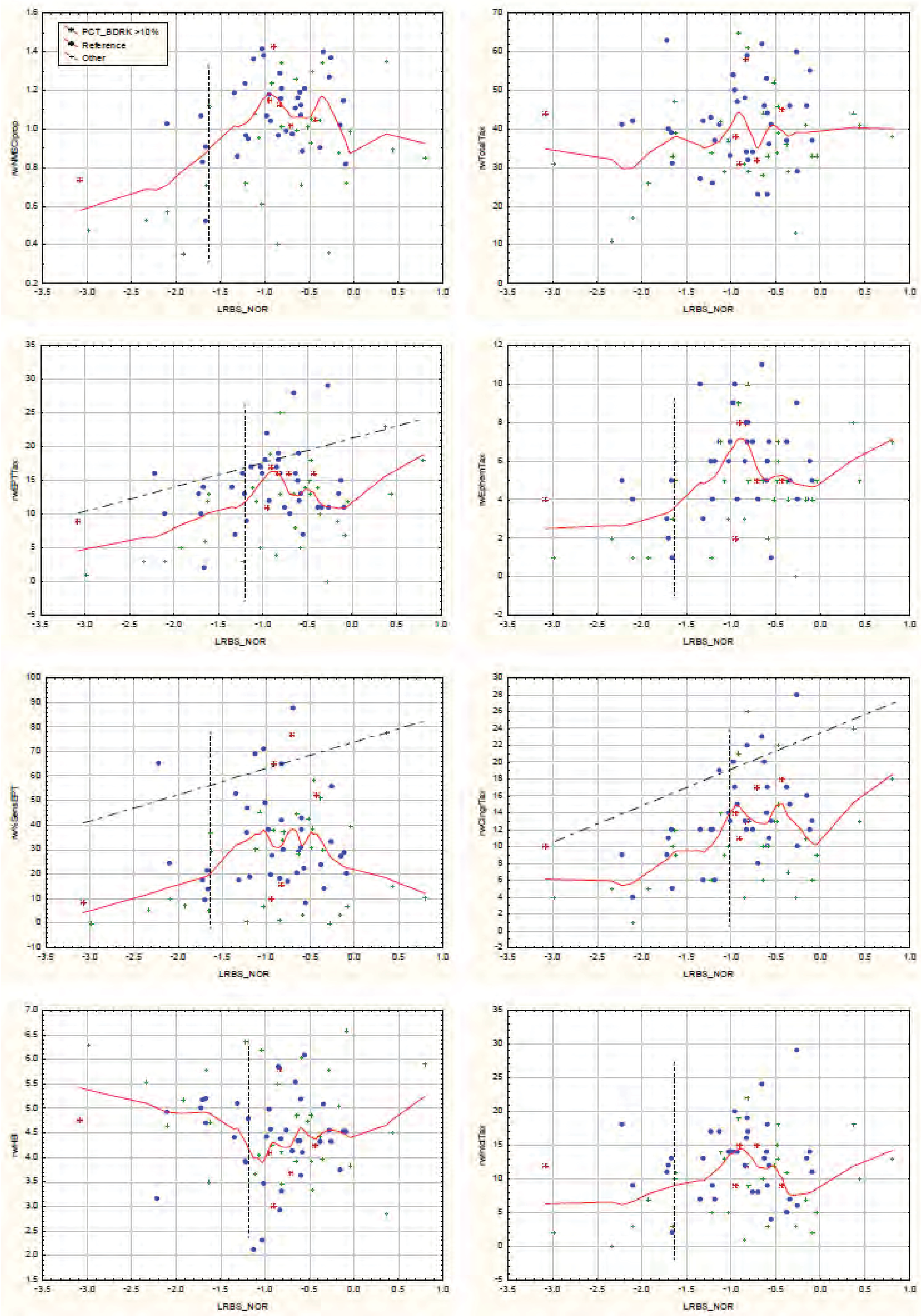

Figure 23. Biological responses to LRBS_NOR in the Mountain site class.

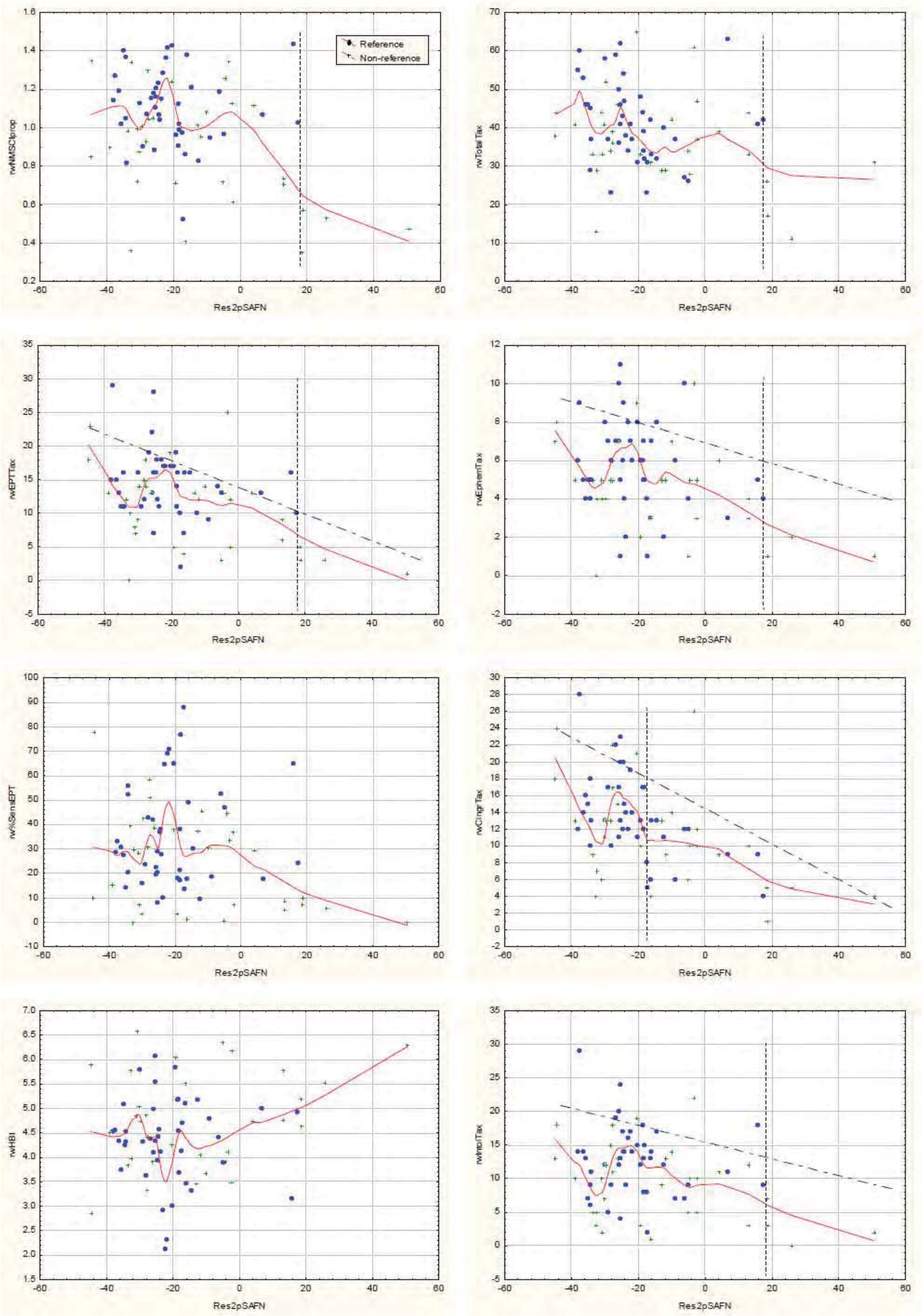

Figure 24. Biological responses to residual % sand & fines in the Mountain site class.

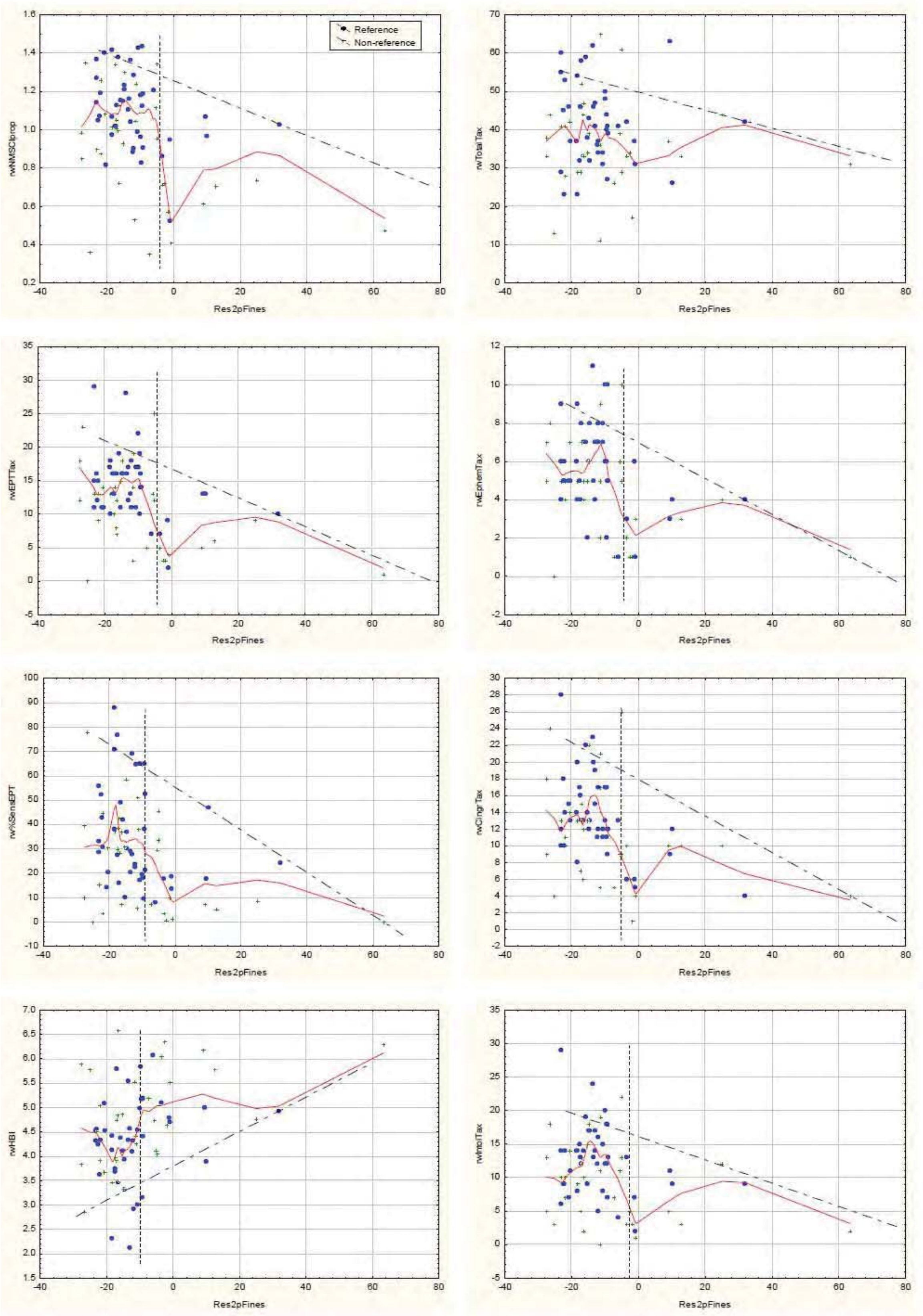

Figure 25. Biological responses to residual % fines in the Mountain site class.

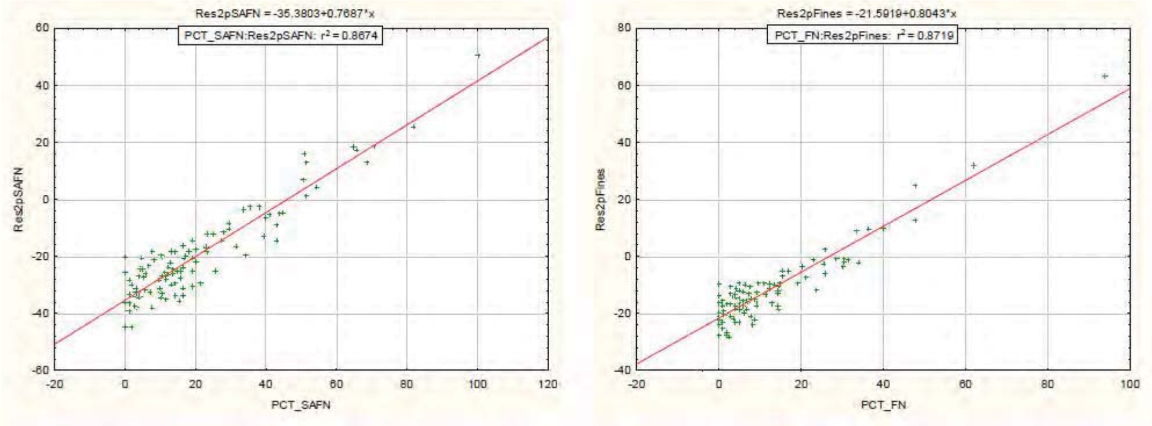

Figure 26. Comparison of untransformed and residualized values for % fines and % sand & fines, in the Mountains.

Foothills

Percent sand & fines (**Figure 27**): Seven of the eight metrics had meaningful upper quantile regressions, indicating that the biological condition is limited by increasing percentages of sand & fine sediments. In five of the eight metrics, a significant change-point was identified in the range of 54-72%. This seems conceptually high and most of the reference sites had values below 40%, ranging up to 61%. The LOWESS regression lines showed changing trends at about 50% sand & fines. Based on the 90th quantile regression line, an increase in 20% sand & fines results in a loss of two taxa.

Percent fines (**Figure 28**): All of the metrics had meaningful upper quantile regressions, indicating that the biological condition is limited by increasing percentages of fine sediments. In five of the eight metrics, a significant change-point was identified between 20 and 32%. This coincides with the LOWESS regression shifts, which start at about 20%, even with those metrics with change-points at higher levels. All reference sites had less than 32% fines. Based on the 90th quantile regression line, an increase in 20% fines results in a loss of one taxon or two EPT taxa.

LRBS (**Figure 29**): Seven of the eight metrics had meaningful upper quantile regressions, indicating that the biological condition is limited by decreasing bed stability. In three metrics, a significant change-point was identified near -2 units. The highest change-point was identified with the clinger taxa metric, which was near -1.8 units. The change-points generally coincide with the LOWESS regression descent as it passes the point of the lowest trough in the higher range. All reference sites have LRBS values greater than -2 units. Based on the 90th quantile regression line, a decrease in 0.5 LRBS units results in a loss of three taxa.

LRBS NOR (**Figure 30**): As with the LRBS measure that includes bedrock and hardpan, seven of the eight metrics had meaningful upper quantile regressions, indicating that the biological condition is limited by decreasing bed stability. Change-points were identified for four metrics, ranging from -1.9 to -1.1. The change-points generally coincide with the LOWESS regression descent as it passes the point of the lowest trough in the higher range. All reference sites have

LRBS values greater than -2 units. Based on the 90th quantile regression line, a decrease in 0.5 LRBS units results in a loss of two-three taxa.

Residual % sand & fines (**Figure 31**): Five of the eight metrics had meaningful upper quantile regressions, indicating that the biological condition is limited by sand & fines in relation to site-specific expectations. In five of the eight metrics, the change-point was identified between two and nine units. The change-points are reasonably associated with the LOWESS regression line trends, which show changes at about 0 units, but the descent does not dip below other troughs until about five units, in most cases. Most of the reference sites had residual values less than 0, ranging up to 8 units. Based on the 90th quantile regression line, an increase in 20 residual sand & fine units results in a loss of two taxa.

Residual % fines (**Figure 32**): All the metrics had meaningful upper quantile regressions, indicating that the biological condition is limited by increasing percentages of fine sediments in relation to site-specific expectations. Significant change-points were identified over a broad range from -12 to 29 units, being lowest for the % sensitive EPTs and the NMMSCI. The lower change-points generally coincided with the LOWESS regression lines, which showed shifting trends at about -12 to -8. The higher change-points are far beyond the highest values in reference sites and are associated with minor changes as indicated by the LOWESS regression lines. Most of the reference sites had residual values less than -8. Based on the 90th quantile regression line, an increase in 20 residual fine units results in a loss of two taxa.

Recommendations:

Based on these analyses of stressor-response in the Foothills site class, the following values and ranges of values are recommended as potential benchmarks for preventing benthic macroinvertebrate degradation.

Percent sand & fines: At greater than 50-60% sand & fines, biological potential is limited. This is conceptually high. The quantile regressions and reference distributions suggest that effects could be occurring at lower levels, but not at any specific level that could be recommended.

Percent fines: Change-points indicate a change at 20-32% fines. To be conservative about protecting degradation of any metric and also in consideration of the high % sand & fines benchmarks, the recommended benchmark is 22% fines.

Log Relative Bed Stability: There seems to be an obvious threshold at -2 units, which is the lowest extent of the values in reference sites. If the % sensitive EPT metric was weighted more heavily than others, the benchmark should be at -1.8 units.

Log Relative Bed Stability NOR: A threshold appears at between -1.9 and -1.1 units, based on significant change-points. The midpoint of this range and a reasonable benchmark is -1.5 units.

Residual of % sand & fines: Based on the median of significant change-points, the recommended benchmark for residual of % sand & fines is 5 units. The regression of % sand & fines with the residual values shows that residual sand & fines of 5 units agrees with 55% sand & fines (**Figure 33**).

Residual of % fines: Based on the effects on sensitive EPTs and the NMMSCI, the recommended threshold for residual of % fines is -12 units. The regression of % fines with the residual values shows that residual fines of -12 units agrees with 20% fines (**Figure 33**).

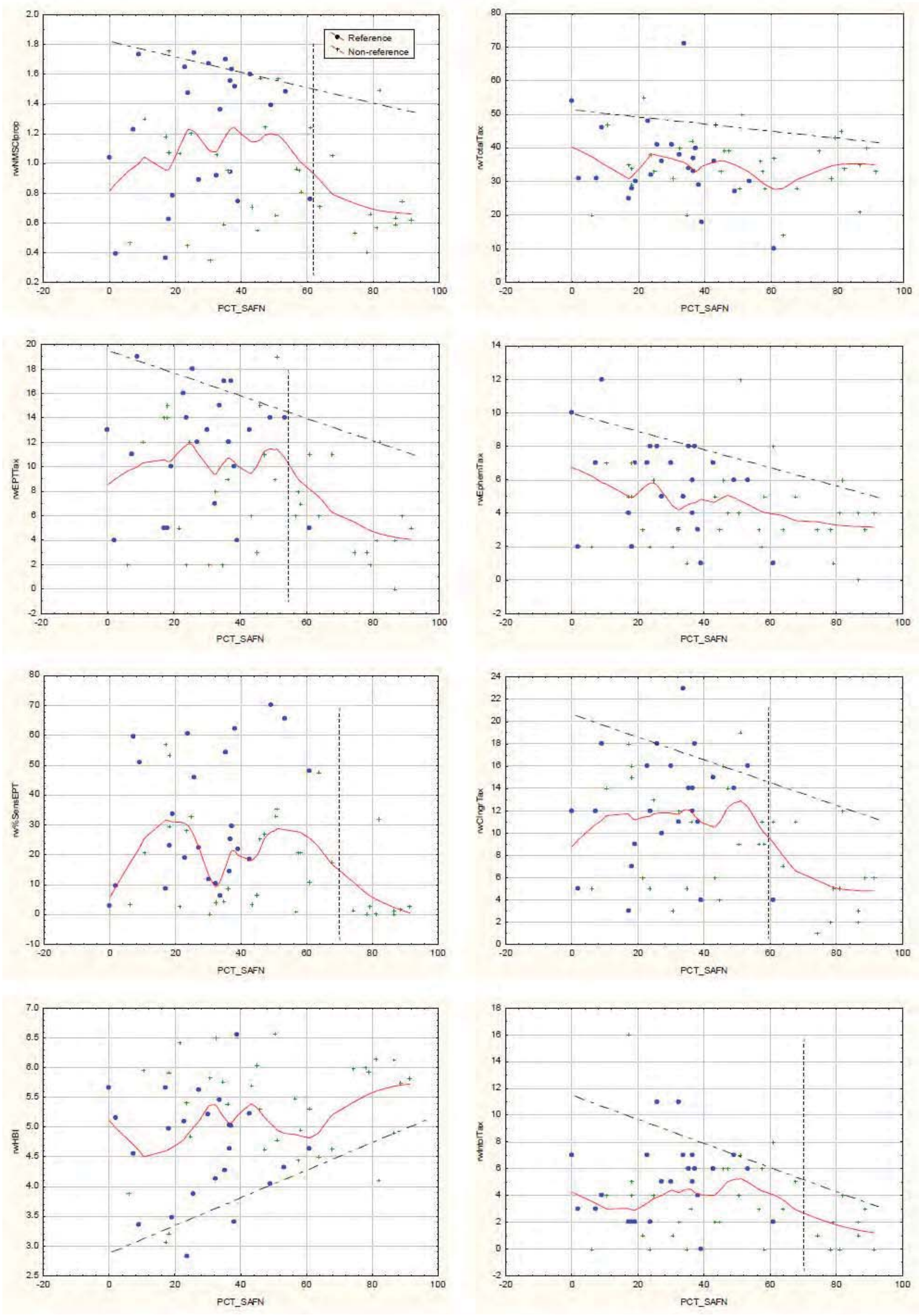

Figure 27. Biological responses to % sand & fines in the Foothills site class.

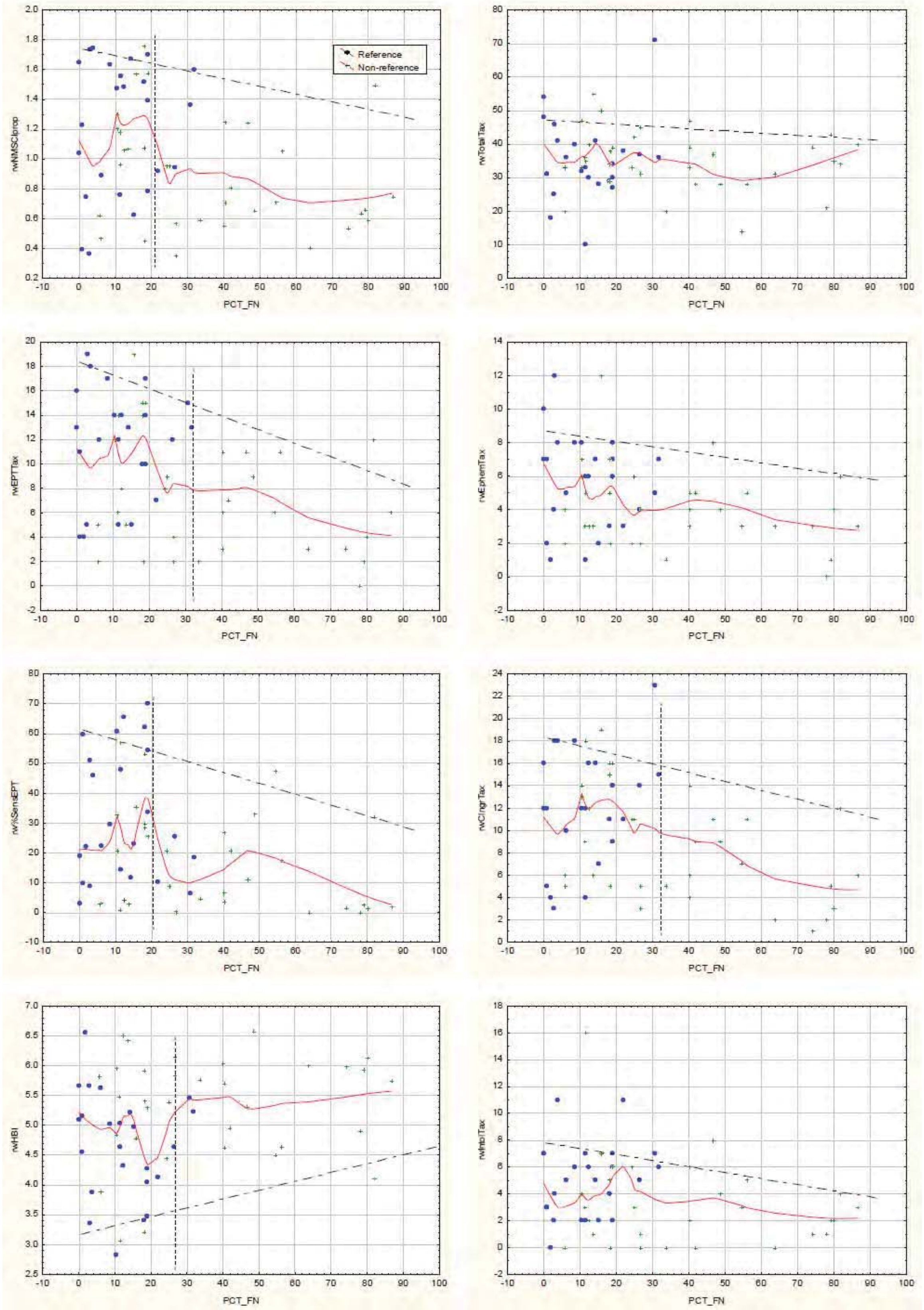

Figure 28. Biological responses to % fines in the Foothills site class.

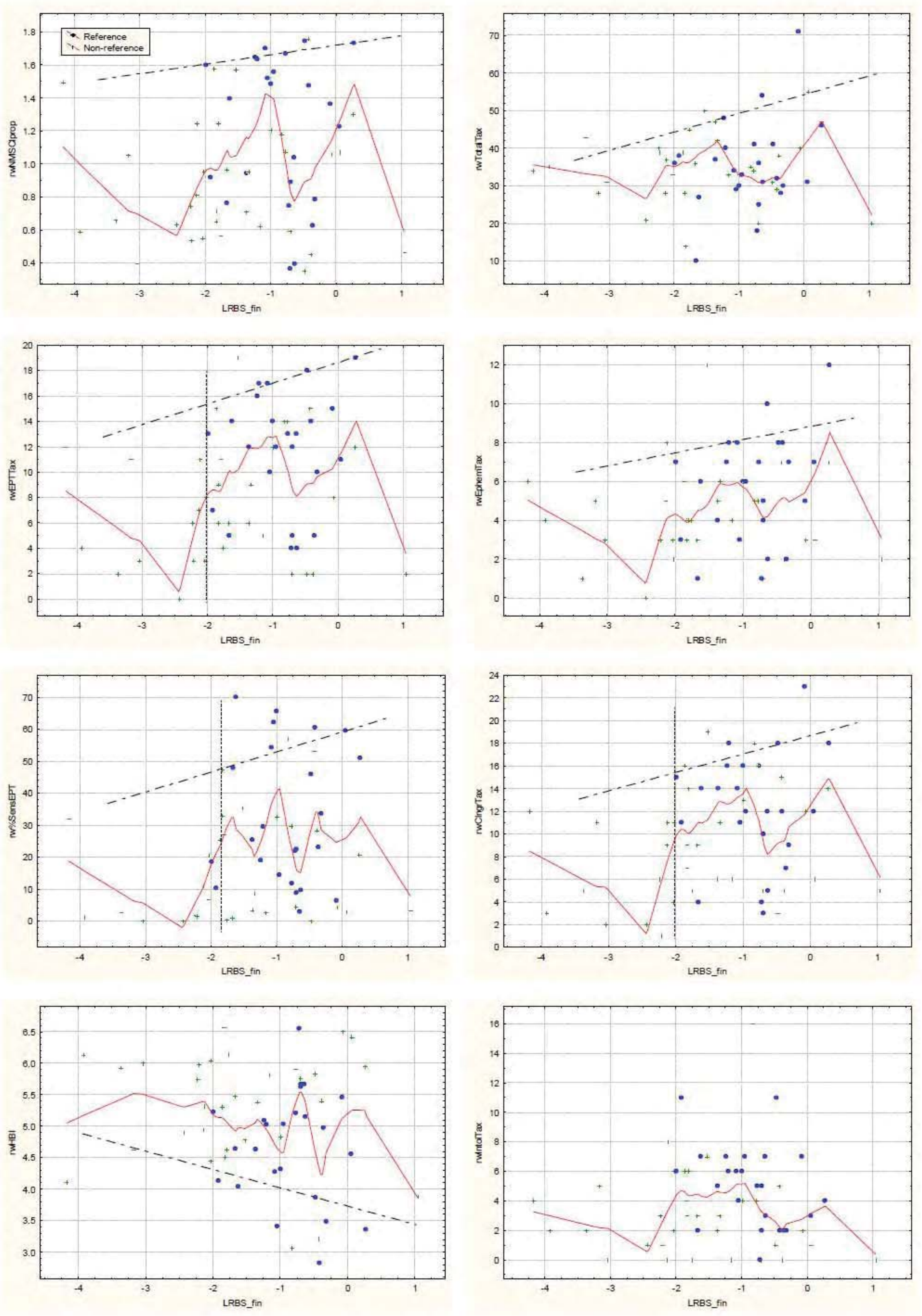

Figure 29. Biological responses to LRBS in the Foothills site class.

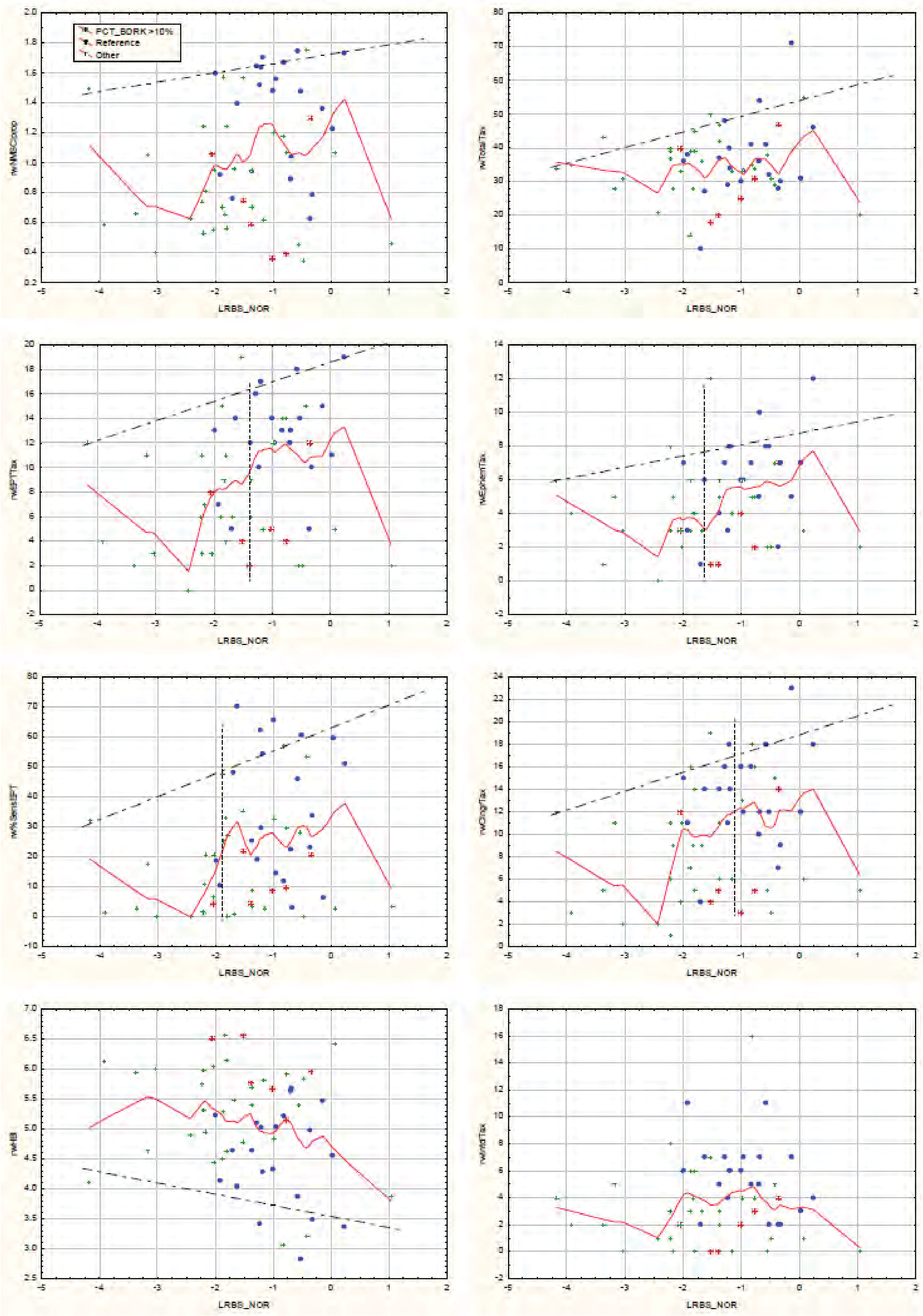

Figure 30. Biological responses to LRBS_NOR in the Foothills site class.

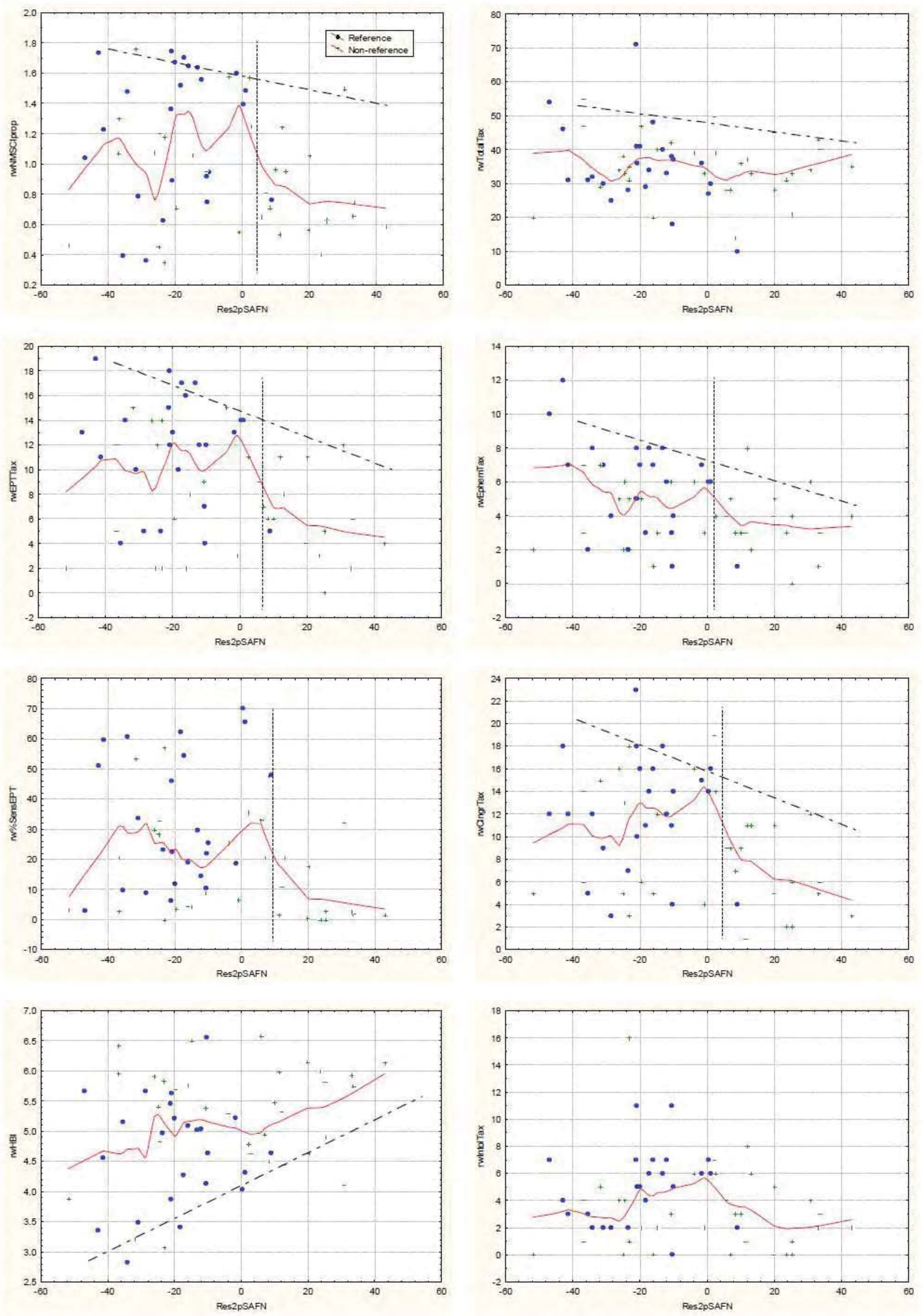

Figure 31. Biological responses to residual % sand & fines in the Foothills site class.

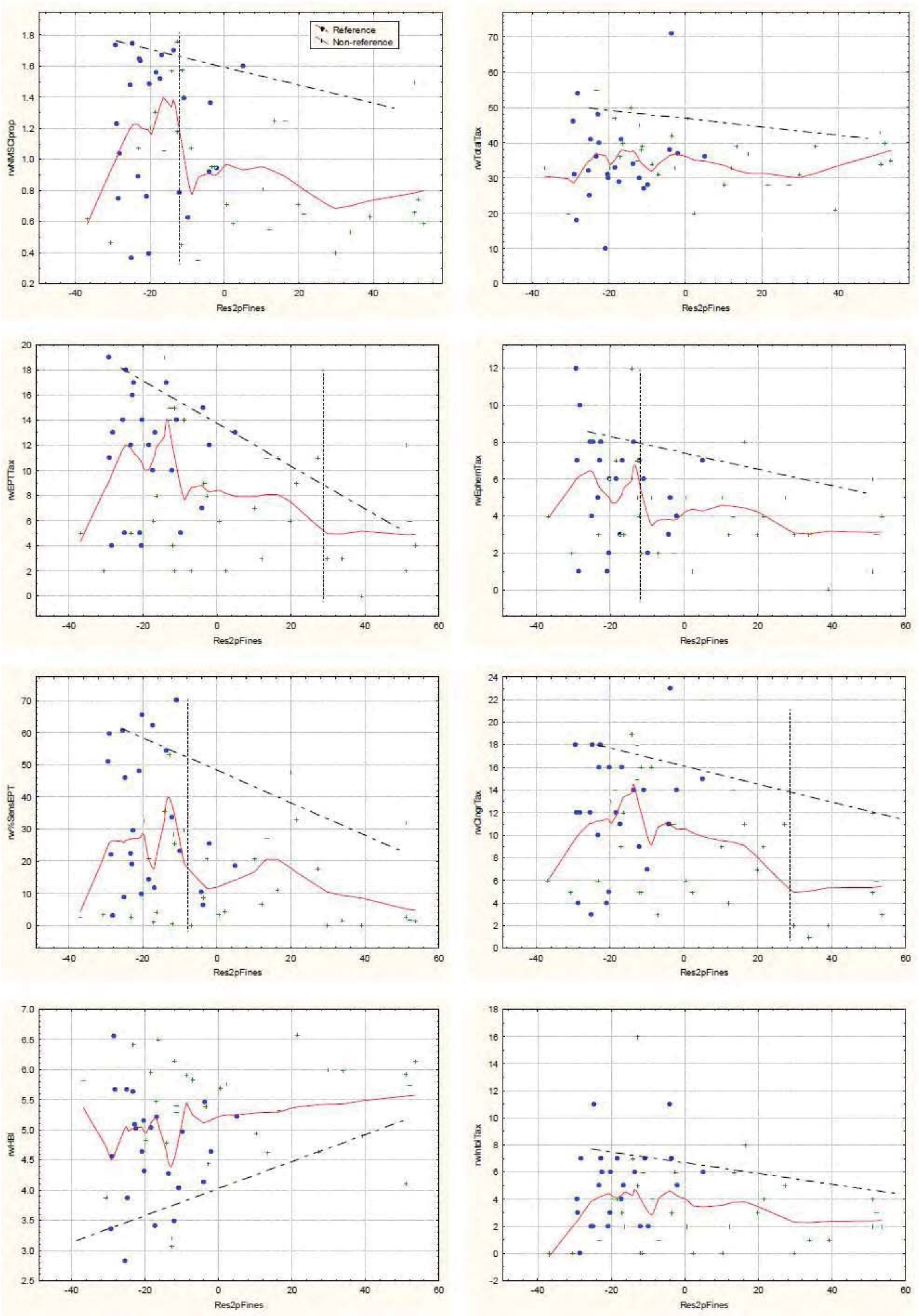

Figure 32. Biological responses to residual % fines in the Foothills site class.

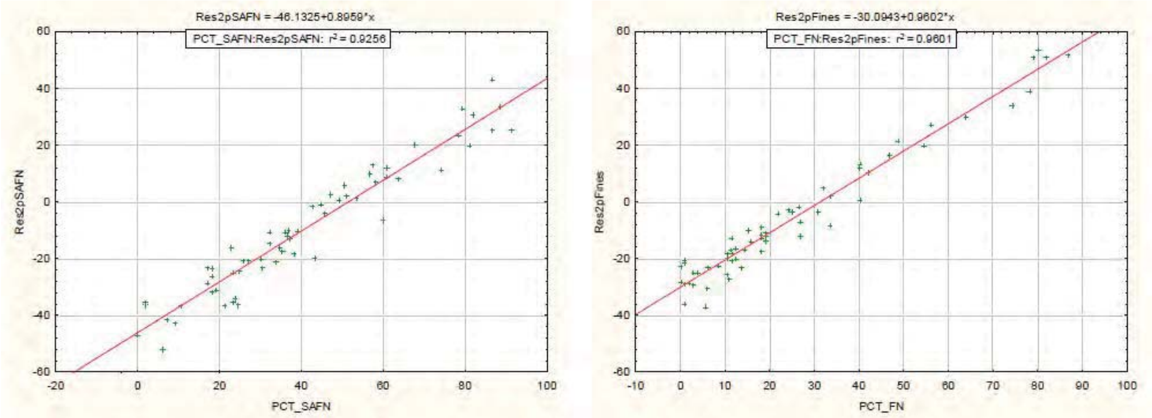

Figure 33. Comparison of untransformed and residualized values for % fines and % sand & fines, in the Foothills.

Xeric areas

Percent sand & fines (**Figure 34**): All of the metrics had meaningful upper quantile regressions, indicating that the biological condition is limited by increasing percentages of sand & fine sediments. The range of change-points was extremely broad, ranging from 28-99%. Five of the eight significant change-point values were between 72 and 75%. The LOWESS regression lines did not show stark or consistent changes in trends. They were more or less gradual, except that EPT taxa and intolerant taxa had very level regression lines up until 74%. Reference sites had values that ranged from 20 to 99%. Based on the 90th quantile regression line, an increase in 20% sand & fines results in a loss of five taxa.

Percent fines (**Figure 35**): Six of the eight metrics had meaningful upper quantile regressions, indicating that the biological condition is limited by increasing percentages of fine sediments. In five of the eight metrics, a significant change-point was identified at about 29%. This generally coincides with the LOWESS regression shifts, which have a peak at about 24%. Reference sites had a broad range of fines (5-87%). Based on the 90th quantile regression line, an increase in 20% fines results in a loss of two EPT taxa.

LRBS (**Figure 36**): All of the metrics had meaningful upper quantile regressions, indicating that the biological condition is limited by decreasing bed stability. Significant change-points were identified in the range of -2.1 to -1.0 units, with half of the change-points occurring at -1.0 units. The change-points generally coincide with the LOWESS regression descents. Several reference sites have LRBS values below -1.0 units and this change-point is lower than that indicated in the Foothills region. Both of these facts suggest caution in identifying a suitable benchmark for LRBS in the Xeric areas (which may relate back to the threshold for the Foothills). Based on the 90th quantile regression line, a decrease in 0.5 LRBS units results in a loss of three taxa.

LRBS NOR (**Figure 37**): The LRBS adjusted to exclude the immobile bedrock and hardpan shows patterns that are similar to the inclusive LRBS. All of the metrics had meaningful upper quantile regressions and significant change-points. Most of the significant change-points separated a few very good metric values from the remaining poorer values at LRBS_NOR values between -1.0 and -1.25 units. The lowest change-point (for the HBI) was at -2.25. The change-points generally coincide with the LOWESS regression descents. Several reference sites have LRBS values below -2.0 units, suggesting that higher benchmarks may be inappropriate. Based on the 90th quantile regression line, a decrease in 0.5 LRBS NOR units results in a loss of three taxa.

Residual Percent Sand & Fines (**Figure 38**): All of the metrics had meaningful upper quantile regressions, indicating that the biological condition is limited by sand & fines in relation to site-specific expectations. In seven of the eight metrics, the change-point was identified between -12 and -14 units. The change-point for total taxa was -1 unit. The change-points are reasonably associated with the LOWESS regression line trends. Reference sites had residual values that spanned the range of all values. Based on the 90th quantile regression line, an increase in 20 residual sand & fine units results in a loss of five taxa.

Residual Percent Fines (**Figure 39**): Four of the eight metrics had meaningful upper quantile regressions, indicating that the biological condition is limited by increasing percentages of fine sediments in relation to site-specific expectations, but not for all aspects of the benthic assemblage. Only two significant change-points were identified at -4 and 5 units, for Ephemeroptera taxa and % sensitive EPTs, respectively. These change-points generally coincided with the LOWESS regression lines, which showed high variability over the stressor gradient. Reference sites had residual values that generally spanned the range of all values. Based on the 90th quantile regression line, an increase in 20 residual fine units results in a loss of two Ephemeroptera taxa.

Recommendations:

Based on these analyses of stressor-response in the Xeric site class, the following values and ranges of values are recommended as potential benchmarks for preventing benthic macroinvertebrate degradation.

Percent Sand & Fines: The best indication from our analyses is that at greater than 74% sand & fines, biological potential is limited. This is conceptually high. The quantile regressions suggest that effects could be occurring at lower levels, but reference values do not support lower benchmarks.

Percent Fines: Change-points indicate a change at 29% fines. Reference and non-reference sites have higher values.

Log Relative Bed Stability: There seems to be a potential threshold at -1 unit. However, this should be considered in the context of benchmarks established for the Foothills.

Log Relative Bed Stability NOR: There seems to be a potential threshold near -1.25 units. This should be considered in the context of benchmarks established for the Foothills.

Residual of % Sand & Fines: Based on the median of significant change-points, the recommended benchmark for residual of percent sand & fines is -12 units. The regression of % sand & fines with the residual values shows that residual sand & fines of -12 units agrees with 70% sand & fines (**Figure 40**).

Residual of % Fines: A potential benchmark could be set at -4 to 5 units, but not with great confidence because of the variability and weakness of the results. The regression of % fines with the residual values shows that residual fines of 0 units agrees with 55% fines (**Figure 40**).

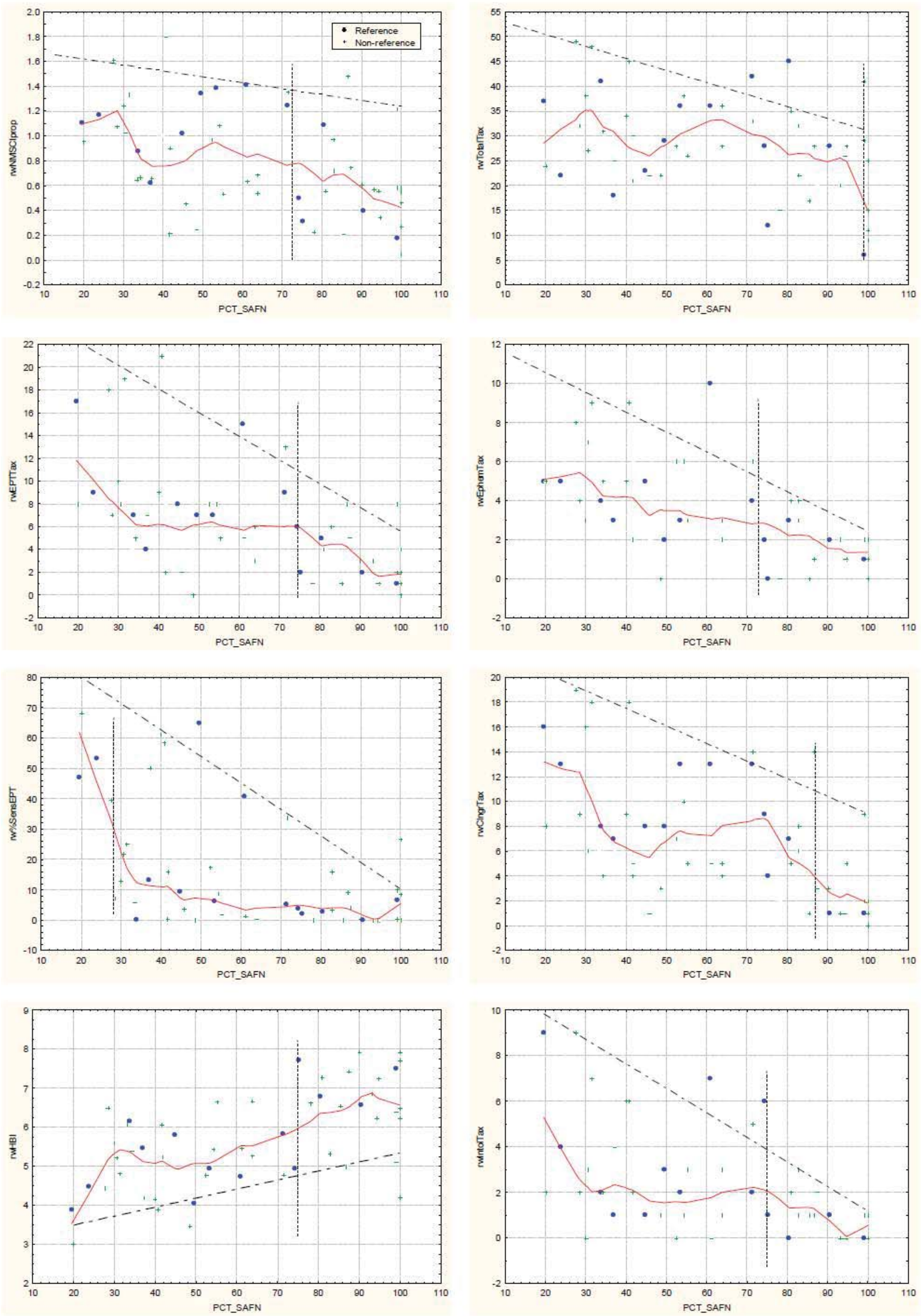

Figure 34. Biological responses to % sand & fines in the Xeric site class.

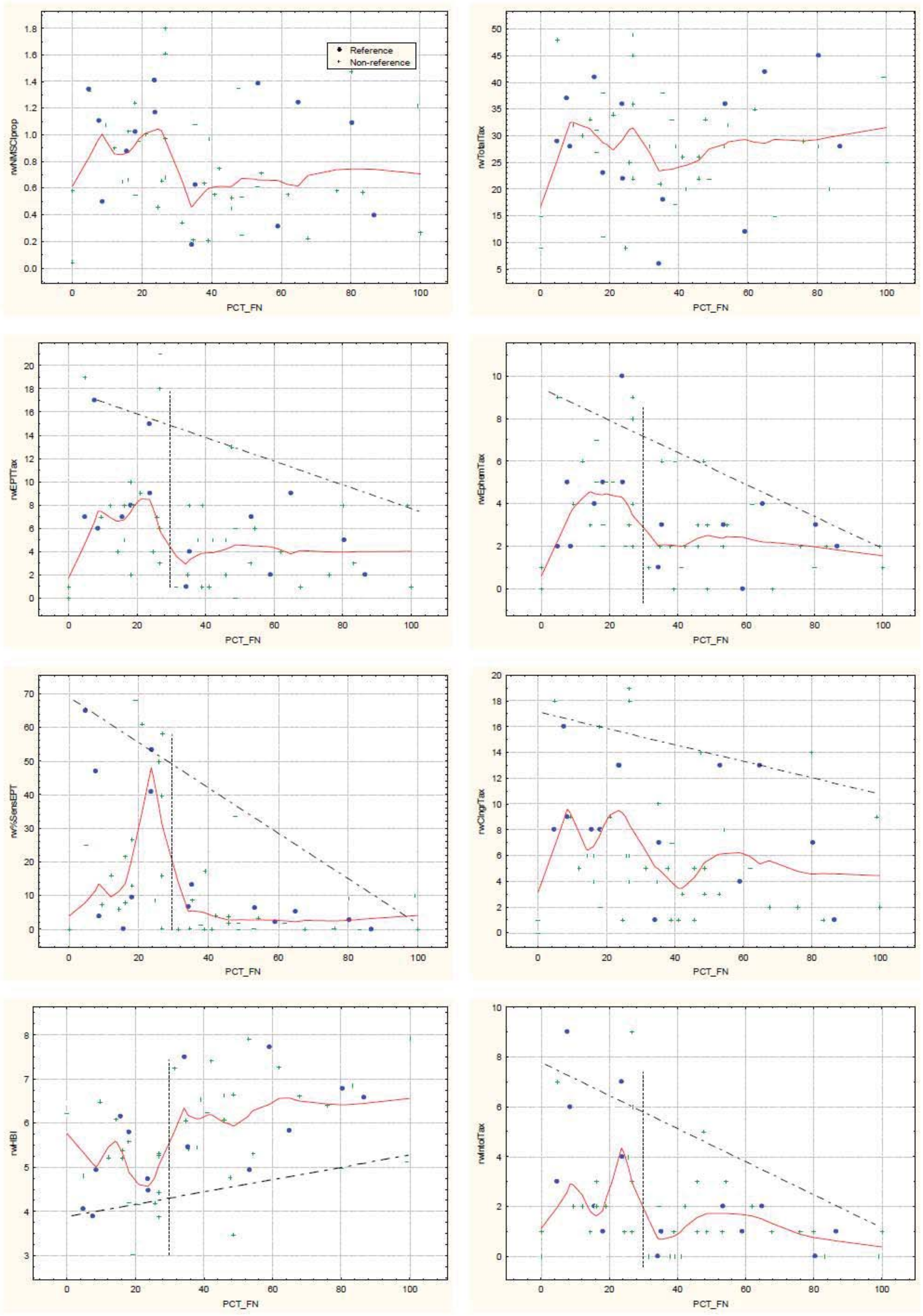

Figure 35. Biological responses to % fines in the Xeric site class.

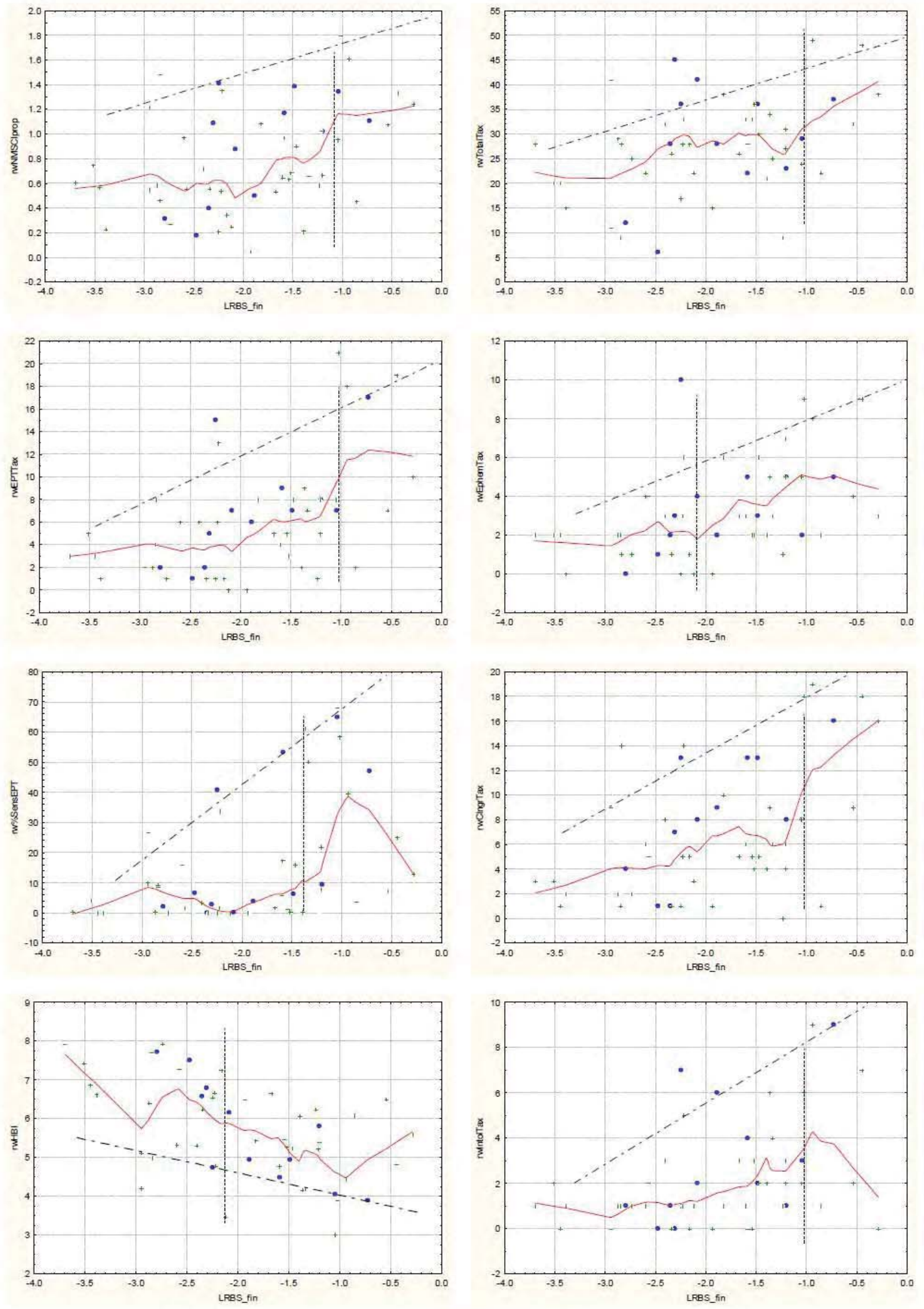

Figure 36. Biological responses to LRBS in the Xeric site class.

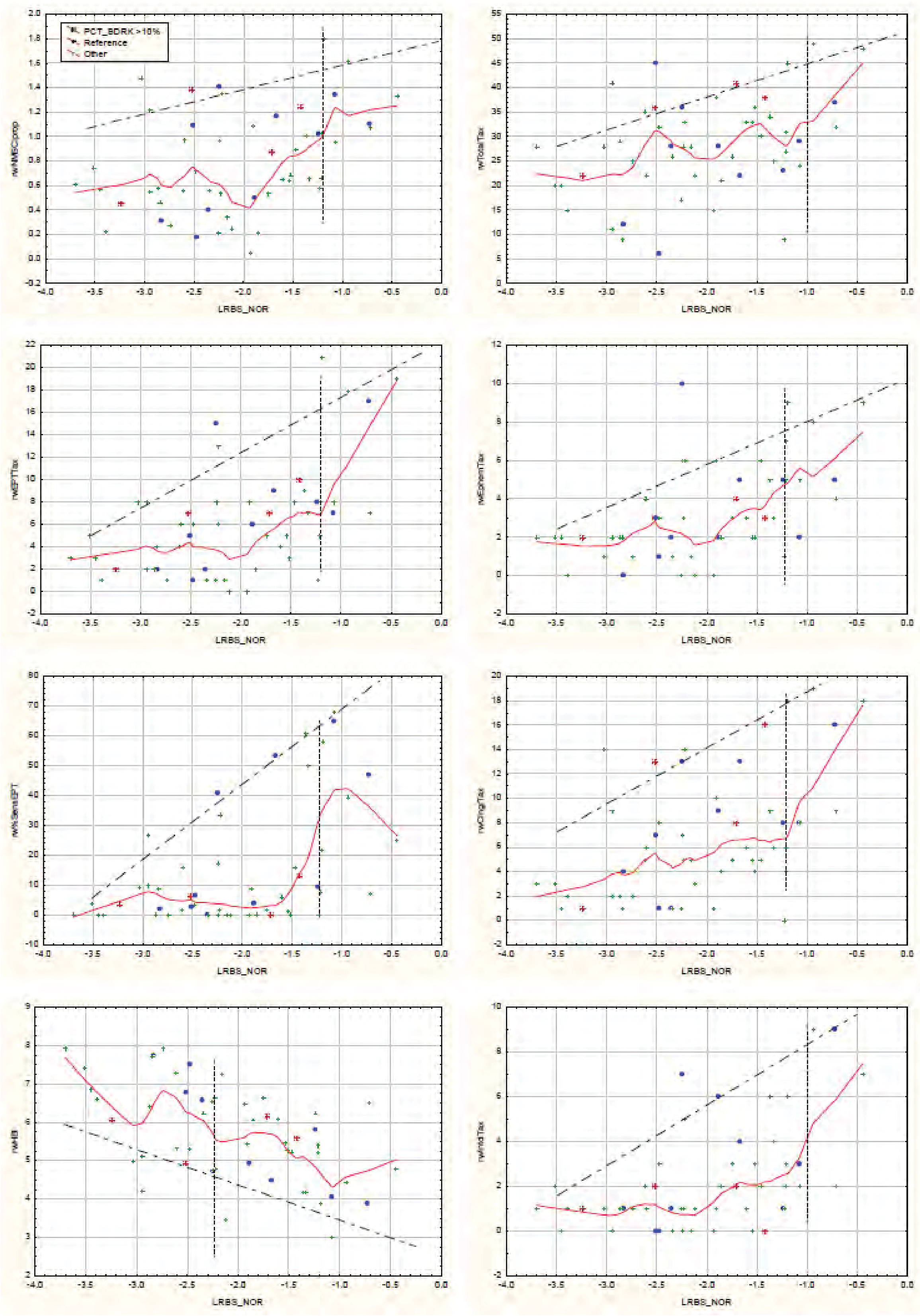

Figure 37. Biological responses to LRBS_NOR in the Xeric site class.

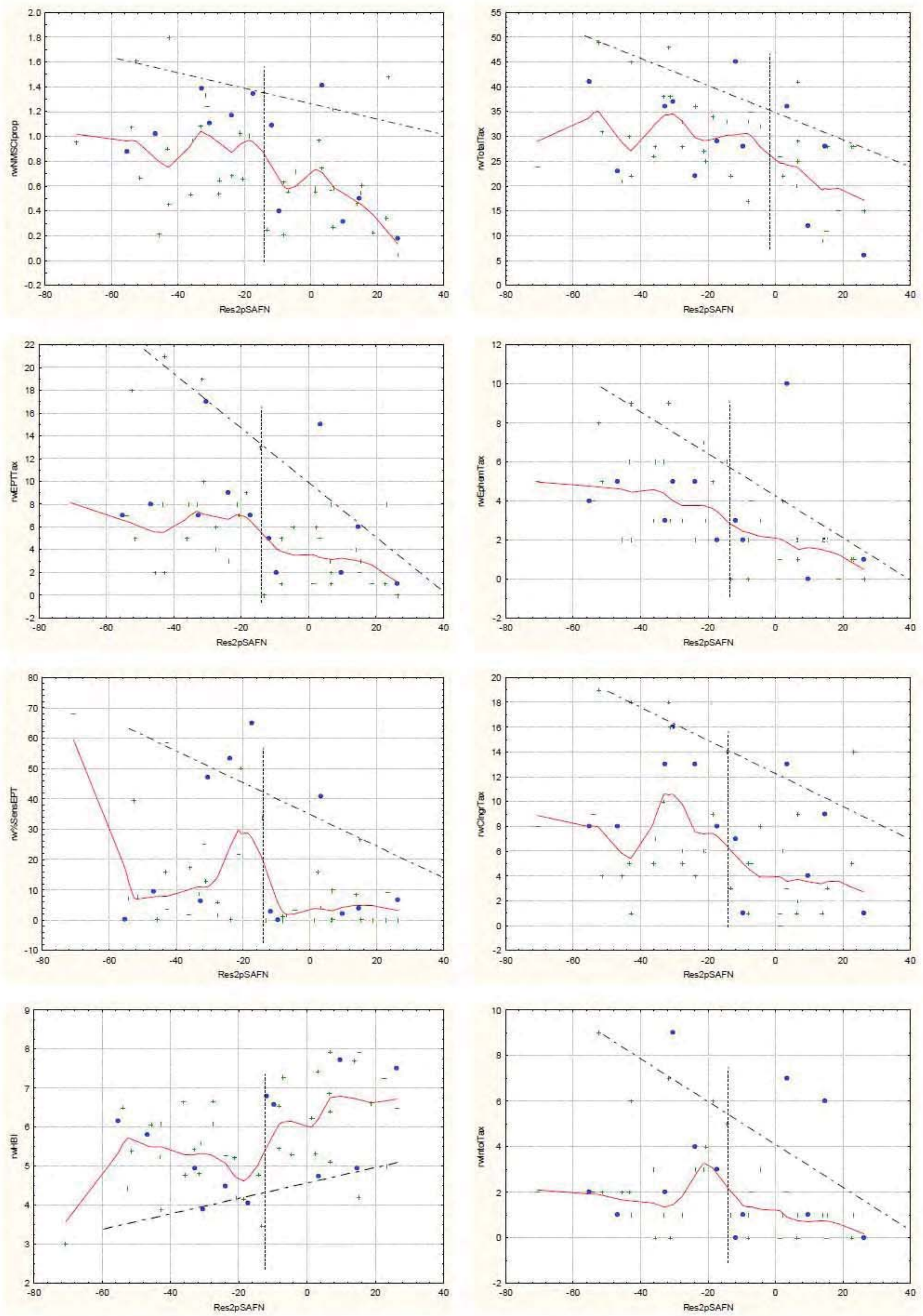

Figure 38. Biological responses to residual % sand & fines in the Xeric site class.

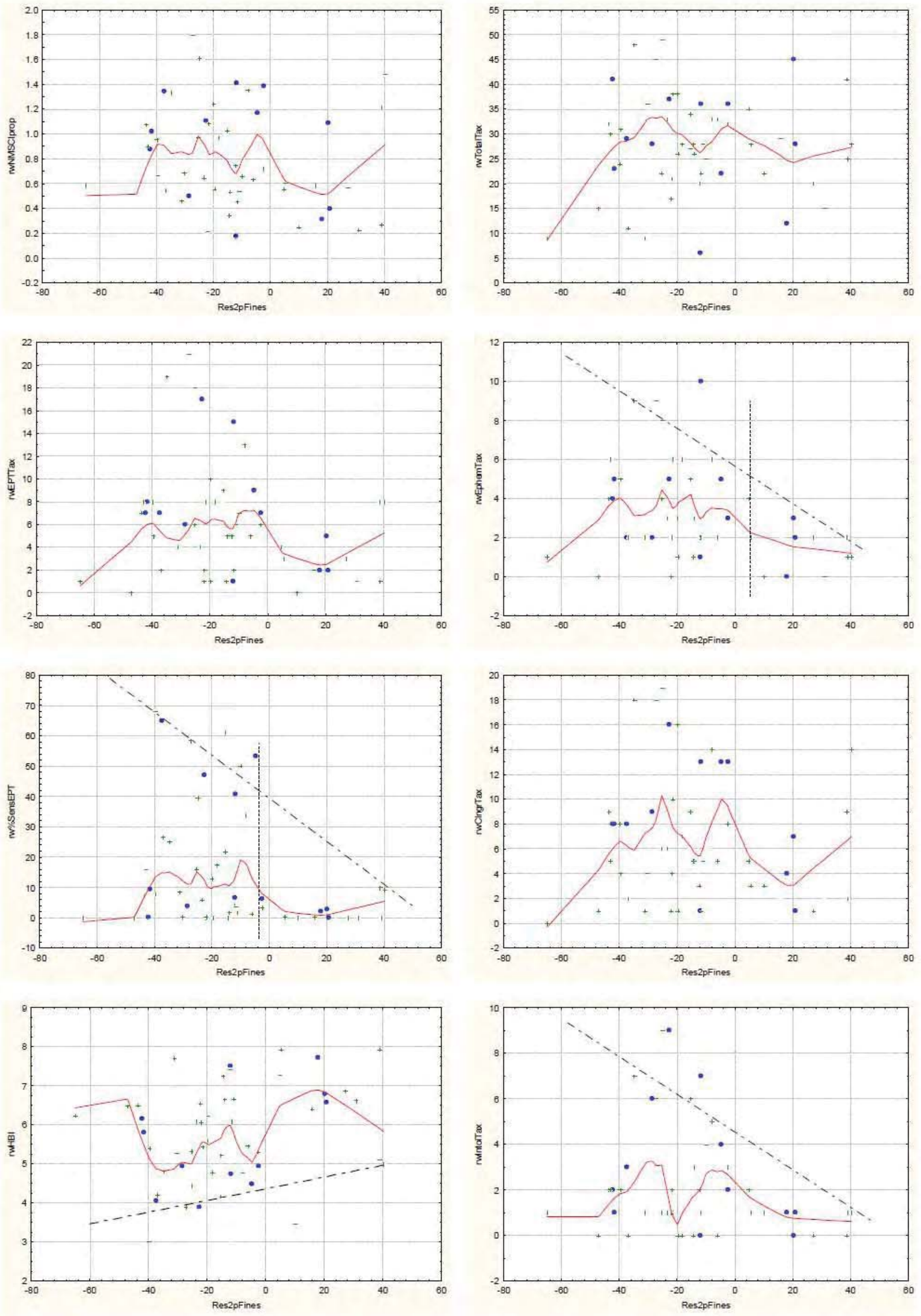

Figure 39. Biological responses to residual % fines in the Xeric site class.

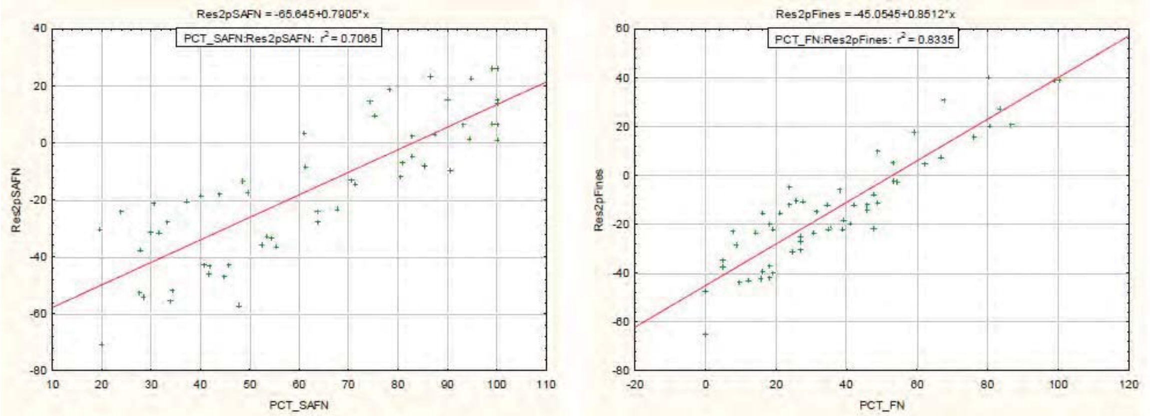

Figure 40. Comparison of untransformed and residualized values for % fines and % sand & fines, in the Xeric areas.

The descriptions of possible effect levels accompanying the graphs above are summarized in **Table 14**. For % sand & fines and % fines, the amount of fine sediments that appears to affect benthic macroinvertebrates is lowest in the mountains and highest in the xeric areas. The residualized values and both forms of the LRBS show variable patterns of benchmarks relative to the site classes so they are not included in the below table.

Table 14. Summary of recommended benchmarks based on biological responses.

Metric \ Site Class	Mountains	Foothills	Xeric
% sand & fines	20%	50-60%	74%
% fines	15%	22%	29%
LRBS	-1.25 units	-1.8 units	-1.0 units
LRBS_NOR	-1.1 units	-1.5 units	-1.25 units
Resid % Sa & Fn	-19 units	5 units	-12 units
Resid % fines	-8 units	-12 units	-4 – 5 units

2.5.2 Suspended sediments and biological responses

Eight metrics were used to assess the effects of suspended sediments during low flow on biological conditions. The NMMSCI was not used because the sites were not categorized by the appropriate classes necessary to apply the NMMSCI. Biological samples used in the analysis were collected with reachwide (RW) methods. Other methods were not included to reduce metric variability that was apparently associated with collection methods in preliminary graphics. Reachwide methods were preferred because the suspended sediments pervade all habitats and the method was common in the database. This analysis includes samples with large and small numbers of individuals.

We concentrated on the log transformed turbidity indicator of suspended sediments. Untransformed distributions were skewed to the low end of the scale and transformation normalized the distributions for easier visualization. Patterns observed with TSS plots (not shown) were similar to those seen in turbidity plots.

We relied on visual interpretation of the bi-plots for assessing possible effects and thresholds. As for the bedded sediment analysis, change-point, quantile regression, and LOWESS regression analyses were performed for the illustrated relationships.

Analyses relating suspended sediment conditions during low flow to biological metrics proved to be ineffective at identifying meaningful thresholds or limiting effects. This may be due to small sample sizes (N = 45-62 in the site classes). In addition, the measurements taken at low flow may not scale up with increasing flows. Thus, streams with relatively high turbidity during low-flows may not be the ones that have biologically stressful turbidity levels during high flows or long durations of higher turbidity. Over the range of turbidity experienced during low flows, turbidity is probably not having meaningful biological effects.

In the Mountains, biological responses were observed in three metrics, Total Taxa, EPT Taxa, and Intolerant Taxa (**Figure 41**). These metrics appear to be somewhat correlated. The best values occur with log turbidity values below 0.2 log units (0.6 ntu). Values gradually drop along the turbidity gradient without any apparent change-points. For four other metrics (% sensitive EPT, HBI, % filterers, % Hydropsychidae), extreme values were observed mid-way in the turbidity scale. Only a single data point had turbidity values greater than 9 ntu (1.0 log [+1] units). Biological metrics for that point showed mixed results, but usually indicated non-reference biological conditions. All Fully Supporting sites in this data subset have turbidity values between 0.35 and 1.0 log units.

In the Foothills, biological associations with log turbidity values were not strong (**Figure 42**). Only one metric, Intolerant Taxa, showed a possible limiting effect of turbidity, where the highest values along the turbidity scale steadily decreased as turbidity increased. The two data points with turbidity values greater than 1.4 log units had mediocre or poor biological metric values. Credible change-points were not apparent.

Biological responses to turbidity are not strong in the Xeric areas (**Figure 43**). Extreme metric values are more common in the middle of the turbidity scale than they are at the ends. LOWESS regression lines appear to show a changing trend at high turbidity values, but that effect appears to be driven by a single data point, which should not carry much interpretive weight. All of the Fully Supporting sites have log turbidity values less than 1.5 units (30.6 ntu). The few points with higher turbidity have mediocre or poor biological conditions.

The generally weak biological response to low-flow turbidity may be because the most stressful high turbidity conditions are not generally observed during low-flow conditions. While we lack duration information with the turbidity data and do not have storm flow turbidity for all sites, we can see from a subset of data that storm-flow turbidities are generally one order of magnitude greater than low-flow turbidity (**Figure 44**). Most of the high outlier sites (in the upper left of the plot) are in the Foothills and Xeric areas. A pattern of site characteristics to explain the highest outliers was not identified, though anecdotally (from GoogleEarth aerial photography) the sites appear to have intense human uses in their surrounding watersheds. This may indicate that the excess suspended sediments in storm flows are from upland sources.

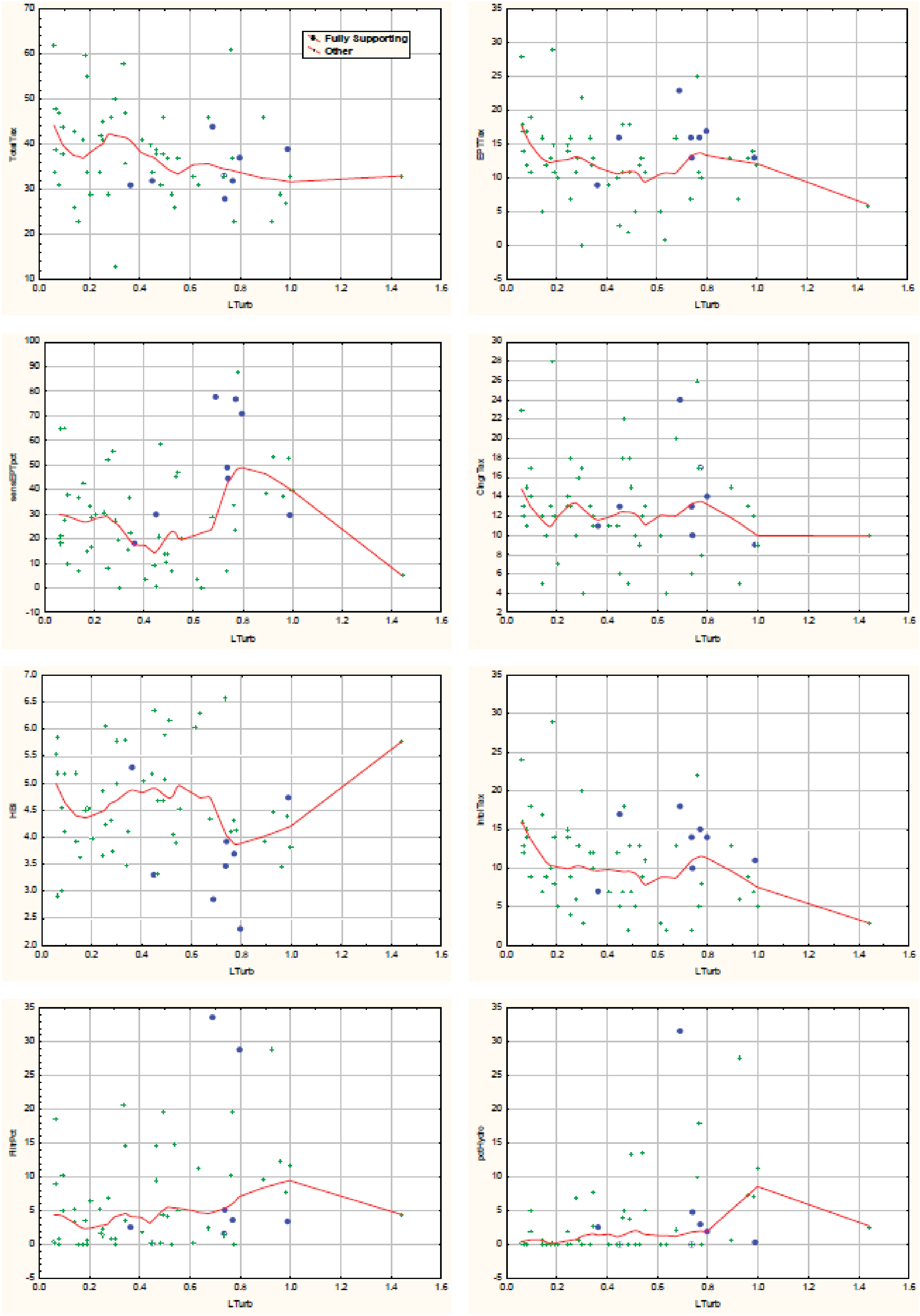

Figure 41. Biological metrics in relation to log turbidity in Mountain sites.

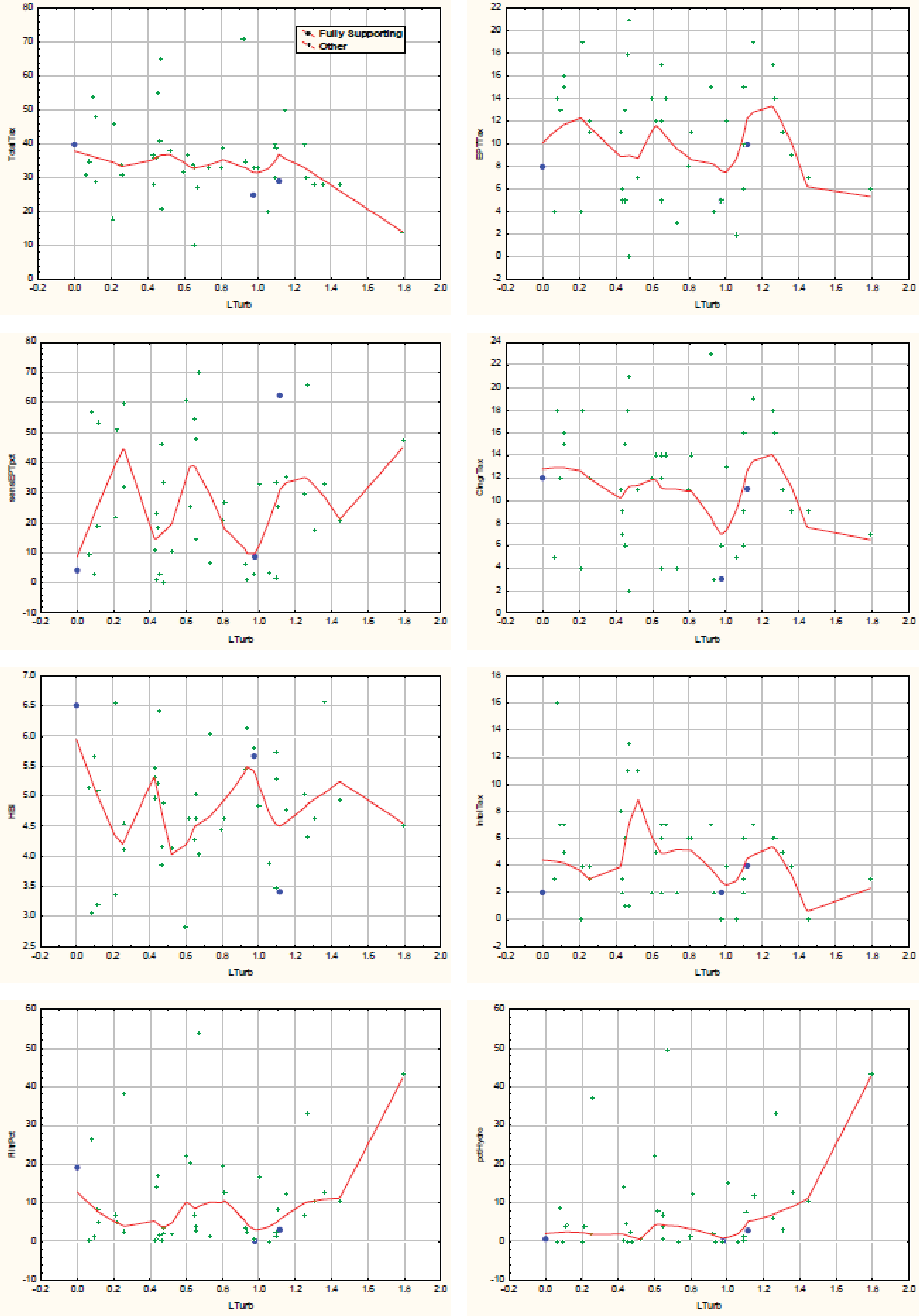

Figure 42. Biological metrics in relation to log turbidity in Foothill sites.

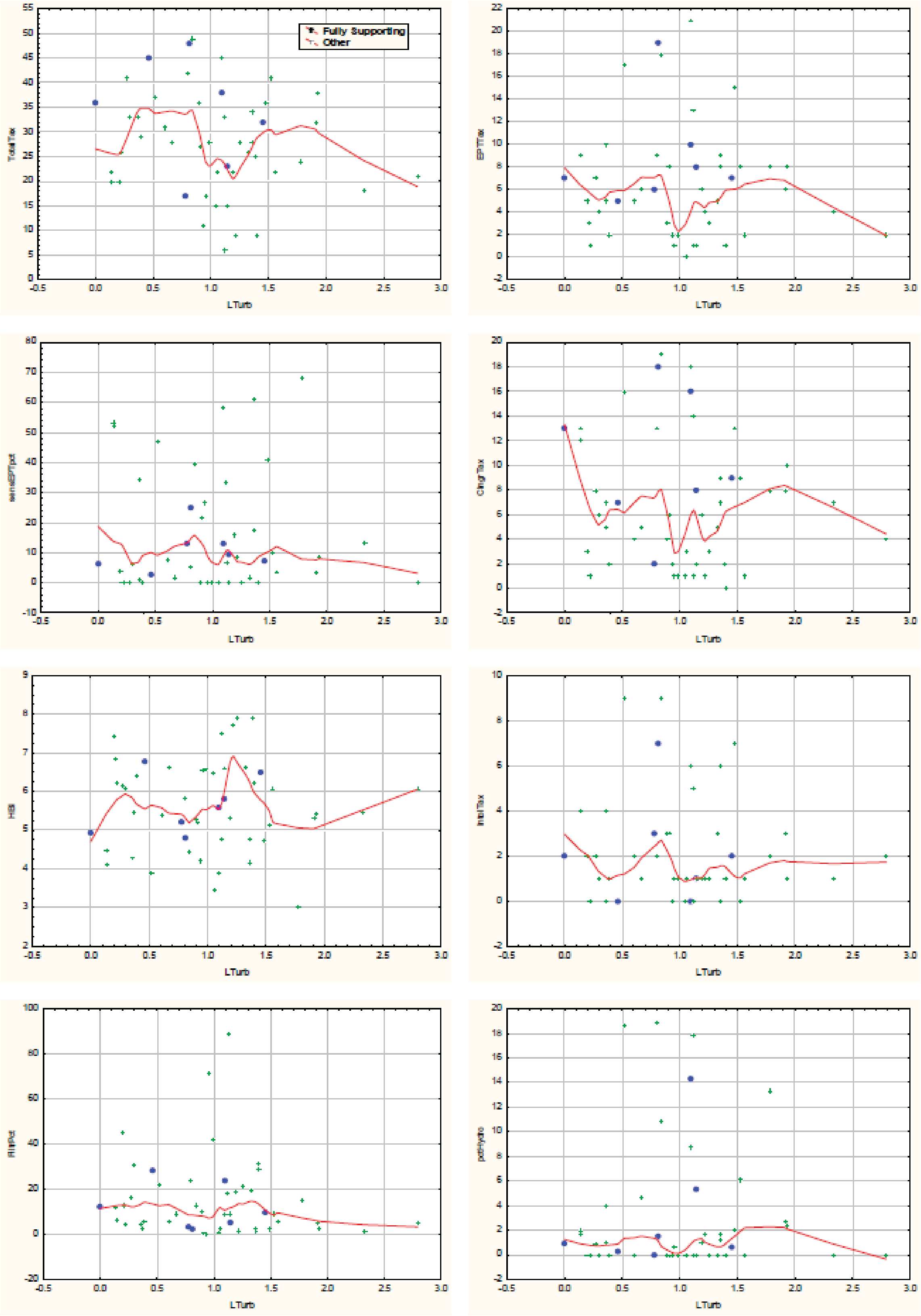

Figure 43. Biological metrics in relation to log turbidity in Xeric sites.

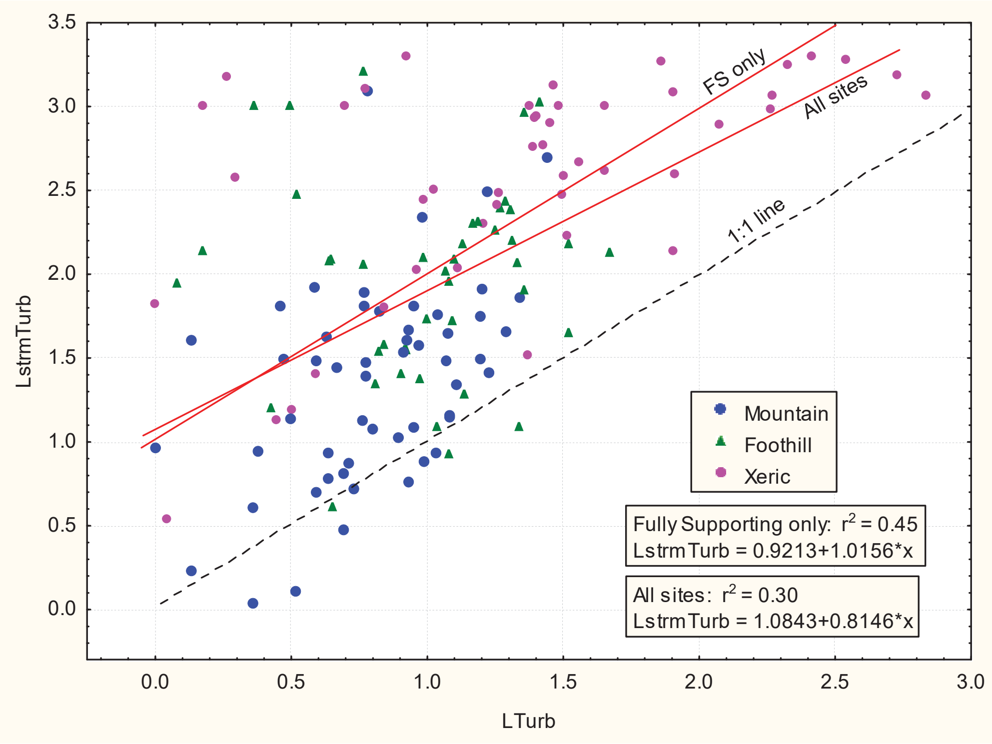

Figure 44. Turbidity compared in low-flow and storm-flow conditions.

Storm flow turbidities that are higher than generally predicted may indicate sediment stresses. The residual of the predicted storm flow turbidity based on fully supporting sites was compared to biological metrics (**Figure 45**). Some of these relationships seem stronger than was seen with the low-flow turbidity, suggesting that the biological effects of turbidity are more important at higher flows. A relatively strong relationship was observed between storm-flow turbidity and intolerant taxa (**Figure 46**). This relationship was also observed in fully supporting sites alone (not shown), highlighting that the storm-flow turbidity values are well-separated by site classes. Relationships between storm-flow turbidity and intolerant taxa were not as strong when calculated in the Mountain and Foothill classes (regression p values = 0.03 and 0.06, respectively). In the Xeric class, the regression was highly significant ($p=0.0001$, $r^2=0.51$).

Therefore, though storm flows are more strongly related to biological measures than low-flow measures, it would be difficult to assign benchmark values to them for two reasons. First, storm-flow measures are limited in this dataset and are variable with flow intensity. Second, natural turbidity expectations seem variable among site classes and sub-setting the data by site class to recognize the natural variability would result in few samples for analysis and weak evidence of any benchmarks.

The workgroup generated a hypothesis that bedded sediment imbalance would be associated with high suspended sediments during storm flows. This could not be examined because insufficient data existed to analyze high-flow suspended sediments and bedded sediments.

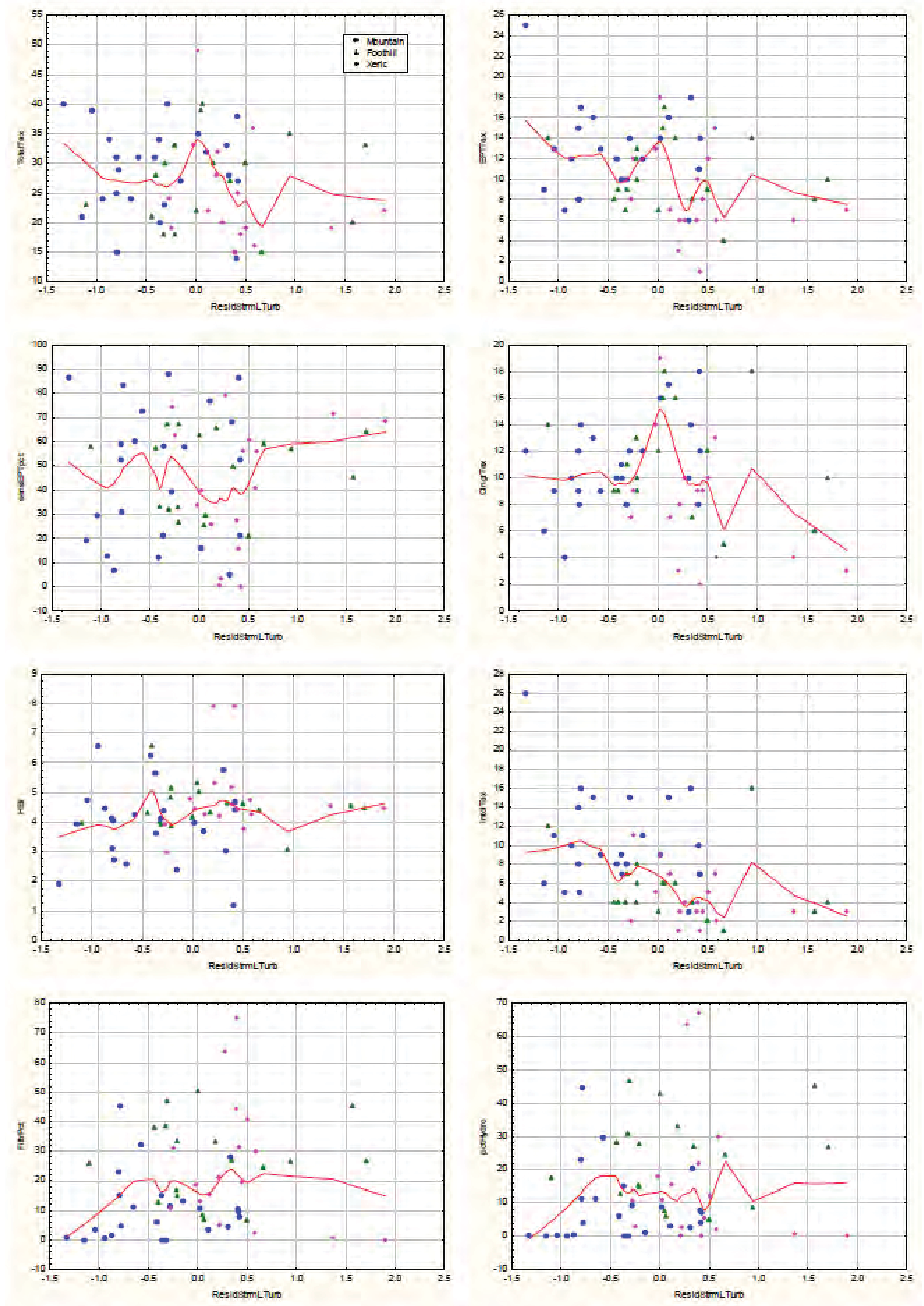

Figure 45. Biological responses to the residual of predicted storm turbidity to observed storm turbidity, in three site classes.

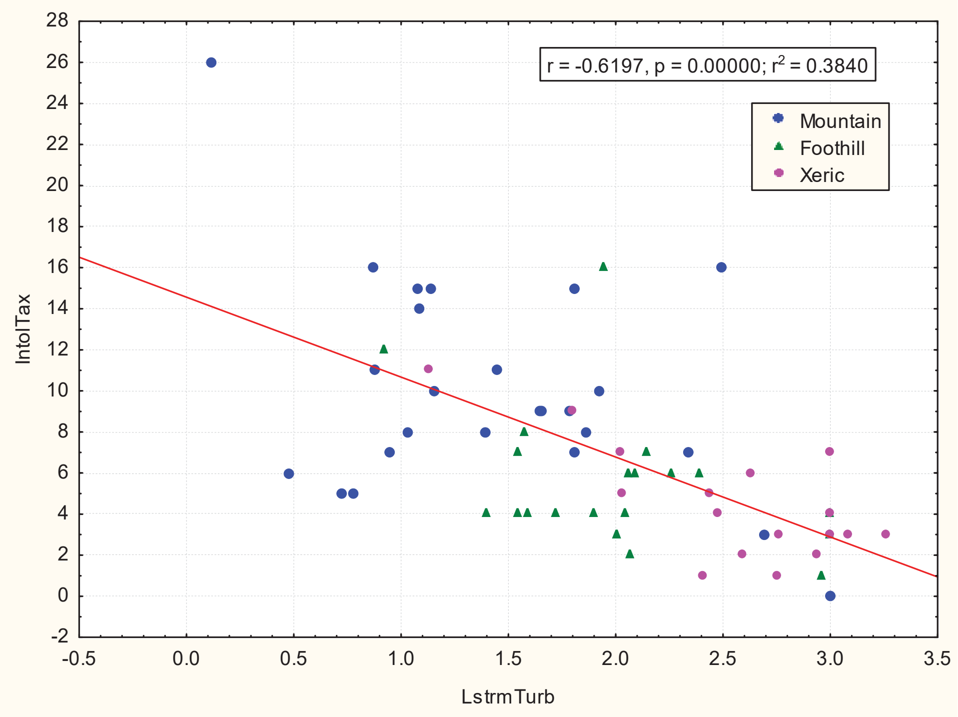

Figure 46. Relationship between the Intolerant Taxa metric and storm-flow turbidity in Mountain, Foothill, and Xeric site classes.

2.6 Potential Sediment Benchmarks and Applications

For each sediment indicator and site class, a benchmark value was recommended through a weight-of-evidence approach that considered the multiple analytical approaches presented above and the strength of each analysis. When formulating a recommendation, considerable weight was afforded the sediment reference condition statistics. These values are the most direct measures of expected sediment conditions in undisturbed sites throughout New Mexico. The sediment reference conditions are characterized independently of the biological conditions. However, they are assumed to exemplify optimal habitat conditions for natural (relatively undisturbed) aquatic fauna.

Corroborating evidence for selection of benchmarks from reference conditions was found in the analysis of associations among sediment and biological indicators. Biological effects are less direct indicators of required sediment conditions because the biota are undoubtedly affected by other environmental conditions, not just sediments.

2.6.1 Bedded sediments

The purpose of the preceding analyses was to provide assessment tools by identifying benchmarks of sediment conditions that can be related to protection or impairment of natural ecological systems, including the biological components. To facilitate both broad, general assessments and detailed evaluations of sediment conditions, a two-tiered procedure is recommended. In the first tier, a measure that is relatively easy to collect and communicate will be compared to a set of known condition benchmarks. This would give a general indication of the sediment conditions, potential biological impairment, and indicate when the second tier is needed. The second tier would give more detailed and different information, allowing refined geomorphic interpretation of the first-tier results.

For bedded sediments, the sediment indicators that are most appropriate for first- and second-tier assessments are % sand & fines and Relative Bed Stability excluding bedrock and hardpan (LRBS_NOR). Percent sand & fines are appropriate because they have precedent in NMED standards, show expected relationships with biological measures, are relatively easy to measure consistently, and are straightforward to communicate. Percent sand & fines may be a better measure than % fines because the sand component has similar modes of biological effect as finer material and fines alone are more variable across sites (and are relatively rare in streams with normal to powerful flows). In Xeric sites, the sand component may be relatively common and the fines may take on more importance.

The LRBS_NOR measure is appropriate as a second-tier indicator because it is scaled to hydro-geomorphic factors of the individual sites, as well as to the broader site classes. This allows evaluation of the potential of the specific site in terms of retaining or flushing fine sediments. When used as a second-tier assessment tool, LRBS_NOR could help explain whether high % sand & fines were expected for a given site or are a result of disturbed conditions. LRBS_NOR can also be used to identify sites with deficient fine sediments, though this condition was not fully explored in the current analysis.

The way that the two indicators can be used for a two-tiered assessment could be as follows for a given site. First, identify the site class and associated benchmarks for % sand & fines and LRBS_NOR. Second, compare the observed % sand & fines to the benchmark and determine whether there is a potential sediment impairment. Then, assuming LRBS_NOR data were collected, compare the observed value to the benchmark to help interpret site-specific sediment conditions. If there is no potential impairment indicated in the first tier, there may not be a need to use the second-tier assessment.

The benchmarks for first- and second-tier assessments can be guided by this report and analysis. However, the final decision on benchmarks is the responsibility of NMED, which may consider factors beyond the scope of this report. Therefore, the following suggestions for benchmark establishment and application are meant only to serve as a starting point.

From the two types of analyses, biological responses and reference distributions, we have a set of summary statistics and benchmarks for the two recommended indicators (**Table 15**). In some cases, as in the Mountain site class, the agreement between the two approaches is remarkably consistent. In the Foothills and Xeric areas, there is a greater difference in benchmarks resulting from the two analytical approaches.

Table 15. Summary of potential benchmarks based on biological responses and reference distributions. Recommended benchmarks are **bolded** and discussed further in the text below.

Metric \ Site Class	Mountains	Foothills	Xeric
Biological Effects			
% sand & fines	**20%**	50-60%	**74%**
% fines	15%	22%	**29%**
LRBS_NOR	**-1.1 units**	-1.5 units	-1.25 units
Reference Distributions			
% sand & fines (reference 75 %ile)	21%	**37% 74%**	
% sand & fines (reference 90 %ile)	35%	45%	84%
% fines (reference 75 %ile)	13%	19%	59%
% fines (reference 90 %ile)	25%	24%	72%
LRBS_NOR (reference 25 %ile)	-1.1 units	**-1.3 units**	**-2.5 units**
LRBS_NOR (reference 10 %ile)	-1.5 units	-1.7 units	-2.7 units

The corroborating analyses in the Mountains yields recommended benchmarks of **20% sand & fines** and **-1.1 LRBS_NOR units**. The protectiveness of these benchmarks can be seen in the relationships of the indicators and the biological metrics (**Figures 20** and **23**). At the recommended benchmarks or near them, biological conditions are generally good on one side and beginning to worsen on the other. This is evident in the better metric values (poor metric values may exist with better sediment conditions due to some unmeasured stressor at those sites) and in the LOWESS regression line. In the analytical dataset, 67% of the Mountain sites had % sand & fines less than 20%, including 71% of reference sites and 50% of stressed sites. Seventy-six percent of sites had LRBS_NOR values greater than -1.1 units, including 75% of reference sites and 67% of stressed sites.

Alternative benchmarks for consideration in the Mountain site class are suggested by some of the metrics with higher change-points (**Figures 20** and **23**) and alternative percentiles of the reference sites (**Table 15**). The higher change-points for % sand & fines range from 50-70% and the 90th percentile of reference data is 35%. For LRBS_NOR, the lower change-points are at -1.6 units and the 10th reference percentile is -1.5 units. Recommended and alternative benchmarks are illustrated in **Figure 47**.

If % sand & fines is used in the Tier 1 assessments, observations above 20% are likely impaired and above 35% are almost certainly impaired. Observations below 20% indicate normal or reference bedded sediment conditions. The Tier 2 LRBS_NOR measure can be used to further qualify the tier 1 assessment results.

In the following graphs (**Figures 47 - 49**), observations in the upper right quadrant show impairment using the Tier 1 indicator, but not using the Tier 2 indicator. In these sites, higher percentages of sand & fines may be natural and impairment may not be indicated. Observations in the upper left quadrant could be assessed as impaired with a high degree of confidence. Likewise, sites in the lower right quadrant could be called unimpaired with a high degree of confidence. Sites in the lower left quadrant have low % sand & fines (passing the Tier 1 benchmark) and low LRBS_NOR values (failing the Tier 2 benchmark). If a site was only assessed using Tier 1 measures, it would not show impairment and might not warrant a Tier 2 assessment. If a Tier 2 assessment was available, the site's assessment would be somewhat uncertain, as it would show low % sand & fines, but otherwise unstable substrates.

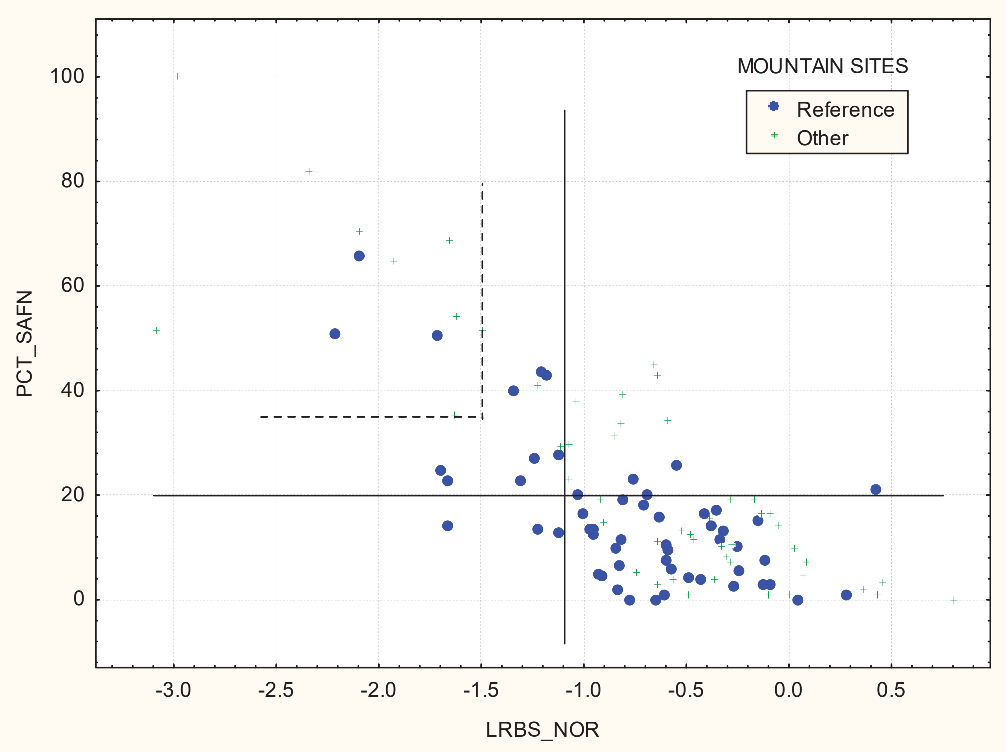

Figure 47. Sediment benchmarks suggested through analyses of biological effects (solid, recommended) and alternative benchmarks based on the 10th/90th reference percentiles (dashed) for the Mountain site class.

The analyses in the Foothills yield recommended benchmarks of **37% sand & fines** and **-1.3 LRBS_NOR units**. These benchmarks are based on the reference percentiles, which are somewhat more protective than the biological change-points. In **Figures 27** and **30**, the best metric values are only on one side of these potential benchmarks, indicating initial effects of bedded sediments. In the analytical dataset, 52% of the Foothill sites had % sand & fines less than 37%, including 74% of reference sites and 44% of stressed sites. Fifty-one percent of sites had LRBS_NOR values greater than -1.3 units, including 78% of reference sites and 44% of stressed sites.

Other potential benchmarks in the Foothill site class are suggested by some of the metrics with higher change-points (**Figures 27** and **30**) and alternative percentiles of the reference sites (**Table 15**). The higher change-points for % sand & fines range from 60-70% and the 90th percentile of reference data is 45%. For LRBS_NOR, the lower change-points are at -1.7 to -1.9 units and the 10th reference percentile is -1.7 units. Potential benchmarks are illustrated in **Figure 48**.

If % sand & fines are used in the Tier 1 Foothills assessments, observations above 37% are potentially impaired and above 45% are almost certainly impaired. Observations below 37% indicate normal or reference bedded sediment conditions.

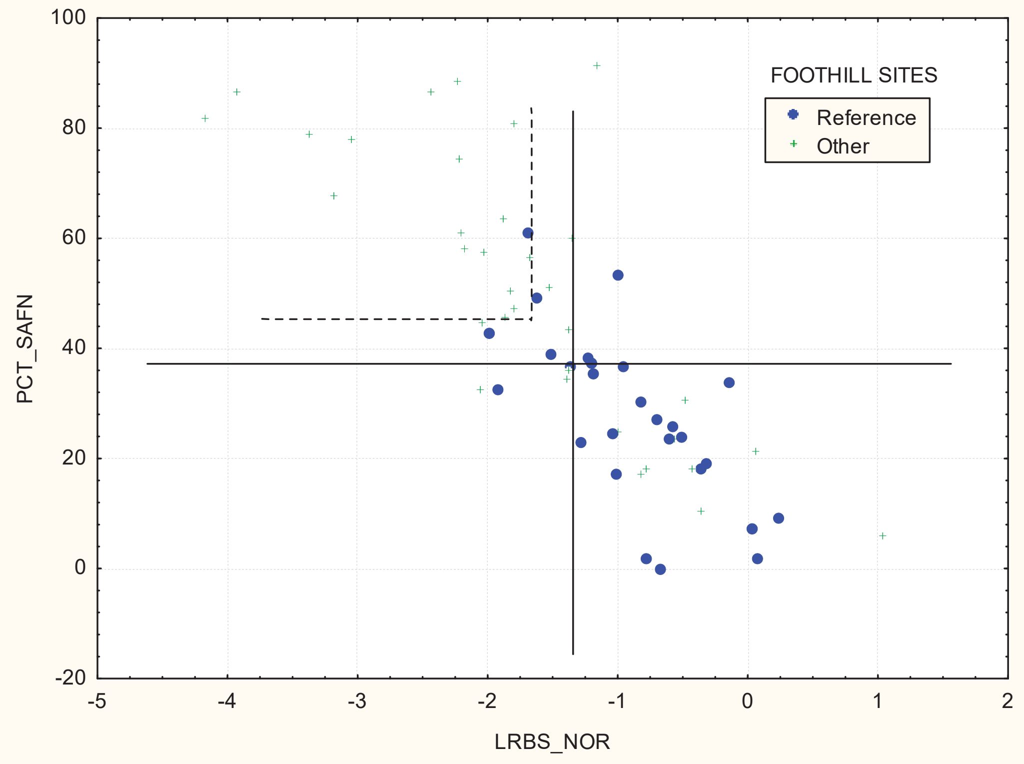

Figure 48. Sediment benchmarks suggested through analyses based on the 25th/75th reference percentiles (solid, recommended) and alternative benchmarks based on biological effects (dashed) for the Foothill site class.

The analyses in the Xeric site class yields recommended benchmarks of **74% sand & fines** and **-2.5 LRBS_NOR units**. The % sand & fines benchmark is based on the corroborated reference percentiles and biological change-points. The LRBS_NOR benchmark would be higher, except that at the level indicated by change-points, more than 71% of reference sites would fail the criterion. Therefore, more weight was given to the reference percentiles, which indicated a benchmark and -2.5 units. In **Figure 34**, the best metric values are generally on one side of the % sand & fines potential benchmark, indicating marked effects of bedded sediments. For LRBS_NOR, the potential benchmark of -2.5 units is at a point beyond which initial biological potential is diminished, and it is limited below the best potential. However, the reference sites with poorer biology still represent undisturbed conditions, and biological response may be attributed to some natural covariate. In the analytical data set, 62% of the Xeric sites had % sand & fines less than 74%, including 71% of reference sites and 40% of the stressed sites. Seventy percent of sites had LRBS_NOR values greater than -2.5 units, including 76% of reference sites and 70% of the stressed sites.

Other potential benchmarks in the Xeric site class are suggested by some of the metrics with more extreme change-points (**Figures 34** and **37**) and alternative percentiles of the reference sites (**Table 15**). The higher change-points for % Sand & Fines range from 87-99% and the 90th percentile of reference data is 84%. For LRBS_NOR, the lowest change-point is -2.25 units and the 10th reference percentile is -2.7 units. Potential benchmarks are illustrated in **Figure 49**.

Sand is a common substrate in the Xeric sites. Therefore, an alternative indicator of sediment stress is % fines, the smaller particle fraction. There is a consistent biological change-point at 29% fines, and this could be applied as an alternative benchmark. However, the distributions of % fines are similar for reference and stressed sites in the Xeric areas and the biological change-point is not corroborated by the percentiles. In the analytical data set, 47% of reference Xeric sites had less than 29% fines, as did 60% of the stressed sites.

If % sand & fines are used in the Tier 1 assessments, observations above 74% are potentially impaired and above 84% are almost certainly impaired. Observations below 74% indicate normal or reference bedded sediment conditions.

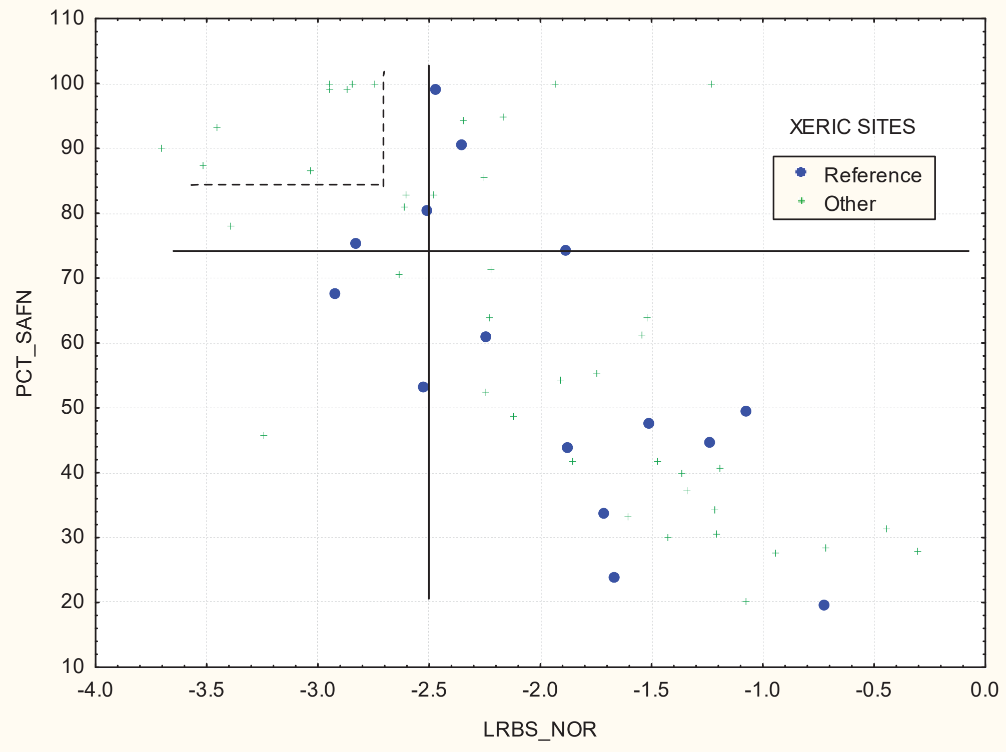

Figure 49. Sediment benchmarks suggested through combined analyses of biological effects and reference percentiles for the Xeric site class. Solid benchmarks are recommended.

Assessment results using the two-tiered system and the recommended benchmarks are presented in **Appendix E**.

2.6.2 Suspended Sediments

Biological metrics were generally unresponsive to low-flow turbidity conditions. The observed non-relationships did not contribute to establishment of turbidity benchmarks. The TSS indicator was highly correlated with turbidity and did not contribute additional information for setting benchmarks from biological responses.

Suspended sediment concentrations during high-flow conditions had a stronger relationship with biological metrics than did low-flow measures. Also, suspended sediments were higher in sites with higher percentages of fine sediments, for both low-flow and high-flow conditions in the mountainous sites (**Figure 14**). These two observations lead to the hypothesis that bedded sediment conditions might be used to predict suspended sediment conditions during high-flows. The hypothesis could not be tested with this data set (few samples in applicable site classes) and benchmarks for suspended sediments could not be established using the relationships between biology, low-flow and high-flow suspended sediments, and bedded sediments.

The remaining technique for establishing suspended sediment benchmarks was through interpretation of the indicator value distributions in fully supporting sites or in all sites. In fully supporting sites, differences in critical percentiles were noted among site classes (**Table 13**). However, within site classes turbidity levels in fully supporting sites were not distinct from levels in other sites (**Figure 15**). Given the lack of biological consequences evident in this data set, the more lenient percentiles (90th) may be appropriate benchmarks until additional data can be collected and analyzed (see **Table 13**). The 90th percentiles in Xeric sites are conceptually high (192 ntu and 263 mg/L), and lower percentiles (such as the 75th) may be more appropriate.

3 Discussions and Conclusions

3.1 Bedded Sediments

Through multiple lines of evidence, we were able to determine bedded sediment conditions that supported benthic macroinvertebrate assemblage integrity and represented least-disturbed reference conditions. The recommended benchmarks are specific to site classes that incorporate much of the natural variability in sediments within the reference data. They can be applied in a two-tier system that allows for rapid or detailed assessments.

Natural variability derived from landscape-scale factors is accounted for by the site classes, which are determined by the level IV ecoregions. We noted that fine sediments were more likely to be high in non-reference sites with erodible lithologies and less likely in any sites with resistant lithologies. Erodibility is related to the landforms and geologies that contribute to ecoregion definition, and to some degree site classes account for erodibility effects. In the assessment approach to be defined by NMED, a qualitative statement about lithologic erodiblility and the likelihood of detecting excessive fine sediments may be appropriate for each site. For those sites with resistant lithologies, assessments that show no sediment impairment would not imply a lack of disturbance at the site or in the catchment.

The recommended benchmarks are quite different among site classes. Expectations for the Mountains are that % sand & fines will be low and LRBS_NOR will be high. In comparison, expectations in the Xeric areas are that fine and mobile sediments are relatively common. Foothills sites should have intermediate sediment conditions. Because of these differences, assigning sites in proper classes is important for attaining accurate assessments.

The benchmarks in **Section 2.6.1** are presented for consideration by NMED. While analyses converge to define the benchmarks presented, there are many benchmark-setting factors that cannot be addressed entirely in this report. For instance, reference and stressed sites are defined with rigor that varies by ecoregion in order to include sufficient samples for analyses. Mountain sites have relatively fewer stressors or stressor sources than Xeric sites. For this reason, the reference percentiles for Mountain sites may be interpreted differently than those in the other regions.

Suggestions for application of the two tiered system are also tentative. While % sand & fines may be easier to measure in the field, return visits to collect two different sets of data may be more difficult than collection of the complete set on the first visit. For that reason, application of both tiers simultaneously may become more common than what is proposed (first apply tier 1 then apply tier 2 only in cases of failed tier 1 assessments). With simultaneous application of the benchmarks for % sand & fines and LRBS_NOR, the lower left quadrant of **Figures 46 - 48** become more interesting. These sites pass the % sand & fines and fail the LRBS_NOR benchmarks. Because the biota seemed less responsive to LRBS_NOR than to % sand & fines, these sites could tentatively be assessed as passing sediment requirements. In **Table 16**, “agreement” is the percentage of site assessments that agree with the reference status, or in other words, the percentage of reference sites passing the benchmark and the percentage of stressed sites failing the benchmark. Agreement percentages near 75% in reference sites reflects use of the 75th or 25th percentile of reference when selecting benchmarks. When applying the two-tiered system with impairment indicated only by failing both indicator benchmarks, the correct identification of reference sites improves relative to each single indicator, while the correct identification of stressed sites is equal or somewhat worse.

Table 16. Agreement of assessment results and reference status for individual indicators and the two-tiered assessment system.

Site Class	% sand & fines benchmark	Agreement % sand & fines Ref/Strs	LRBS_NOR benchmark	Agreement LRBS_NOR Ref/Strs	Agreement 2 tiers Ref/Strs
Mountains	< 20	71/50	> -1.1	75/33	80/33
Foothills	< 37	74/56	> -1.3	78/56	85/44
Xeric	< 74	71/60	> -2.5	76/30	87/30

In the Xeric sites, an alternative benchmark can be considered for % fines, where > 29% would indicate stress. This benchmark is based on biological responses. Reference and stressed sites show poor to fair agreement with this benchmark (47/60), which should therefore be used with caution. Combination of % fines with LRBS_NOR in a two tiered system does not show improved performance when substituted for % sand & fines in this dataset.

Another consideration in applying benchmarks is the distinction in assessments based on alternative benchmarks, as indicated in **Figures 47-49**. A failure of the alternative benchmark is more certainly associated with biological impairment than failure of the recommended benchmark alone. Therefore, failure of either Tier 1 or Tier 2 alternative benchmarks might trigger a failed assessment regardless of agreement among Tier 1 and Tier 2 indicators.

As NMED selects the final benchmarks, Type I and Type II assessment errors should be considered. Type I errors, incorrectly assessing a reference site as impaired, occurs in 13 to 20% of reference sites using the two tiered system in this analytical data set. Type II errors, incorrectly assessing a stressed site as un-impaired, occurs in 56 to 70% of stressed sites. The error appears to be unbalanced, allowing impairment to go undetected in many sites. There are valid reasons for accepting higher Type II than Type I error. A prominent effect that we observed was that disturbance at the site or in the watershed often did not result in higher stream sediments in sites with resistant lithology. Since the stressed sites are not defined using measures that certainly cause sediment stress, we can expect greater Type II error in the sediment indicators. A lack of disturbance, as measured using our reference site criteria, appears to reflect a lack of factors that cause sediment stress. In addition, we have more reference sites than stressed sites. Therefore, we have more confidence in the definition of reference sites and the agreement of the indicators with that definition than we have of stressed sites and agreement in that group.

Benchmarks for the residual values for % fines and % sand & fines could be deduced from the analyses. The residuals were highly correlated with the unadjusted percentages. They are more difficult to calculate and do not appear to add interpretive information or assessment accuracy. Therefore, the raw percentages will probably be more useful in assessments.

3.2 Suspended Sediments

The low-flow turbidity measure was relatively unrelated to benthic macroinvertebrate metrics. However, there were two relationships suggesting that further examination of the critical relationships would be worthwhile. First, low-flow turbidity was positively related to high-flow turbidity. Second, high-flow turbidity was related to benthic macroinvertebrate metrics. Therefore, we expect that low-flow turbidity gives some indication of the high-flow conditions that may have more turbidity (by an order of magnitude), and are apparently more stressful to biota.

While there was some indication that sites with higher percentages of sand and fine sediments yield higher turbidities during low-flow and high-flow conditions, the low-flow relationship was noisy and the high-flow relationship was represented by only a few data points. Neither could be used to predict turbidity from % sand & fines.

We suspect that high-flow turbidity values were variable with specific discharge in each stream. However, we did not have sufficient information to determine the degree to which the storm flows exceeded low-flows, and thus could not establish sediment transport curves relating flow to turbidity. For this reason and others relating to sampling feasibility, it did not make sense to suggest high-flow turbidity benchmarks.

The recommended suspended sediment benchmarks are based entirely on the distribution of turbidity and TSS values in Fully Supporting sites. Fully Supporting streams were used as a surrogate for reference streams in the suspended sediment dataset. As surrogates, the Fully Supporting streams do not appear to represent sites with the best attainable sediment conditions. The lack of detected stressors is not necessarily sufficient for identifying the best attainable conditions. They are designated based on multiple stressors, including suspended sediments in cold-water streams. Including sediment conditions to define sediment reference sites introduces circularity into the argument for defining sediment benchmarks.

In future suspended sediment analyses, definition of reference and stressed sites based on variables that are not direct measures of sediment might improve the evidence of distinctions between reference and stressed sediment conditions. In particular, reference criteria related to upland disturbance and activities might prove useful. In our analysis, we saw that sites with intensive upland activity detected in aerial imagery seemed to have particularly high turbidity in high-flow conditions compared to low-flow conditions.

3.3 Research needs

While the presented analyses inform decisions regarding benchmark selection for sediments in New Mexico streams, they also generate additional questions that could not be thoroughly answered without further effort.

Precision analysis on the indicators would allow NMED to assess measurement error and temporal variability, which would enhance interpretation of sediment assessment certainty. Many of the data points used in this analysis were from single grab samples or observations. Multiple data points were available for each site in a small number of cases, and average values were used in analyses. Same-day replicate sampling was not evident in the data set, but would be necessary to assess measurement error while controlling for temporal changes.

Collection of sufficient data for calculation of LRBS is somewhat cumbersome or time consuming in the field. NMED has expressed interest in estimation of conditions in general categories as an alternative to extensive measurements. Conceptually, estimated channel roughness is feasible, but should be tested before implementation. Testing could include comparison of LRBS calculated from both quantitative measures and estimations, including estimations made by different field crew members for calibration.

Lithologic erodibility appears to control the degree to which disturbance at a site or in the catchment results in excess fine or mobile sediments. The workgroup hypothesized that minor changes in sediments in resistant lithologies might have comparable biological effects as major changes in erodible lithologies. The reasoning was that biota in resistant lithologies are more dependent on stable conditions than biota in streams with commonly shifting streambeds. Because reference streams in resistant and erodible lithologies were indistinguishable as site classes – they had similar sediment characteristics – the sediment sensitivities of biota within the two erodibility groups were not assessed. Information on minor sediment-biota relationships in resistant lithologies would help identification of sediment stress in resistant lithologies, but would require sensitive and precise sediment and biological measures. Lithologic erodibility may be important in suspended sediment analysis also, and such data would be worth developing.

In our analysis, it appeared that biota were more responsive to absolute measures of sediment composition than to measures of sediment relative to the stream potential, like LRBS or the residual measures of % fines and % sand & fines. However, differences in sediment composition occur not only in response to disturbance, but also in the full range of natural settings. The fauna that are naturally adapted to fine and mobile sediments occur with fine sediments, but may not indicate disturbance unless the presence of such sediments are unexpected. The measures used in the recommended two-tiered system include one absolute measure and one relative measure. The requirement that both indicators fail the benchmarks before impairment is assessed assures that the signal at the site is of both regionally abundant fine sediments and sediment amounts that are more than the stream can efficiently transport. Our understanding of the conditions to which biota respond would be enhanced with additional examination of the biotic responses to relative sediment supply along the natural gradient of fine sediment abundance. Obviously, the causes and mechanisms of effects would be difficult to tease apart.

Our uncertainty regarding suspended sediment benchmarks reflects some data limitations as well as confounding effects. The data limitations relate to single samples for each site (in many cases); assumptions regarding low-, runoff-, or storm-flows; and inadequate reference site determination. These limitations could be removed through more intensive suspended sediment sampling, collection of additional ancillary data, and more detailed site characterizations (on-site or remote sensing). Much of the stress associated with suspended sediments is related to duration of exposure as well as concentration (Newcombe and Jensen 1996). While repetitive sampling in one stream may limit the number of streams that can be sampled, the more intensive temporal information taken in a variety of flow conditions, coupled with simultaneous flow information, habitat observations, and biological sampling, would allow complete characterization of the suspended sediment conditions and effects. A concerted effort to collect site information and develop reference site criteria for all suspended sediment sites would allow development site classes and descriptions of reference conditions.

The confounding factors related to suspended sediments include bedded sediment conditions, as these may be sources or sinks for suspendable sediments. While we attempted to related suspended and bedded sediment conditions in this report, the data for such an analysis were sparse because there was no concerted effort, especially with regards to multiple flow regimes at each site.

TSS and turbidity were related over several streams in generally similar streams. Because the two variables measure different components of the suspended materials, the relationships that were developed for the whole data set might not hold true in individual streams. Stream classes could be refined to identify the types of streams that have similar TSS/turbidity relationships, which could be a factor of stream power, lithology or other sediment sources, bedded sediment composition, and hydrologic variability.

Pursuit of high-flow suspended sediment data would be worthwhile, though costs of collection of sufficient data might be prohibitive. A number of streams would need to be sampled in multiple flow conditions to establish sediment transport curves. Then the curves could be related to biological responses. Therefore, streams of reference and stressed sediment quality in all stream site classes would need to be represented in sufficient numbers for analysis.

Of equal importance to suspended sediment concentration as a stress mechanism is suspended sediment duration (Newcomb and Jensen 1996). Longer exposures to high concentrations are more stressful than shorter exposures. Because low-flow conditions are somewhat stable, we assume that the sediment conditions observed during low-flow represent chronic conditions. When multiple suspended sediment readings can be taken over time at individual sites, the duration element of exposures could be addressed. As with sediment transport curve data, collection of sufficient data for analysis may be prohibitively expensive.

Assessment of multiple biological assemblages would enhance interpretations of stressor-response relationships. Fish samples may be obtained as well as macroinvertebrate samples to show different sensitivities among the entire biotic system. Such data were not available for the current effort.

Unbalanced sediments undoubtedly occur in the presence of other stressors. A multiple stressor analysis focused on isolating sediment effects may not be possible without sediment chemistry, but other correlated factors could be assessed. In our analysis, we correlated sediment indicators and biological metrics, which implies a causative relationship that was never proven. Causative analyses may be infeasible, though more effort could be applied towards identifying sites in which multiple stressors are either prominent or lacking.

4 References Cited

In addition to the following publications there are two data sets that were used in analysis and are available as electronic **Appendices F**, the datasets and the data dictionary.

Arizona Department of Environmental Quality (ADEQ). 2006. Draft Technical Support Documentation for the Narrative Biocriteria Standard. ADEQ Water Quality Division, Hydrologic Support and Assessment Section, Surface Water Quality Standards & Monitoring Unit, Phoenix, AZ.

Bryce, S.A., G.A. Lomnicky, P.R. Kaufmann, L.S. McAllister, and T.L. Ernst. 2008. Development of Biologically Based Sediment Criteria in Mountain Streams of the Western United States. North American Journal of Fisheries Management 28:1714–1724

Bryce, S.A., G.A. Lomnicky, and P.R. Kaufmann. 2010. Protecting sediment-sensitive aquatic species in mountain streams through the application of biologically based streambed sediment criteria. Journal of the North American Benthological Society, 2010, 29(2):657–672.

Cade, B. S., J. W. Terrell, and R. L. Schroeder. 1999. Estimating effects of limiting factors with regression quantiles. Ecology 80(1):311-323

Chapman, D.W. and K.P. McLeod. 1987. Development of Criteria for Fine Sediment in Northern Rockies Ecoregion. United States Environment Protection Agency, Water Division, Report 910/9-87-162, Seattle, Washington, USA.

Cleveland, W. S. 1979. Robust locally weighted regression and smoothing scatterplots. Journal of American Statistical Association 74: 829-836.

Cormier, S.M., J.F. Paul, R.L. Spehar, P. Shaw-Allen, W.J. Berry, and G.W. Suter II. 2008. Using Field Data and Weight of Evidence to Develop Water Quality Criteria. Integrated Environmental Assessment and Management. 4(4):490–504.

Jacobi, G.Z., M.D. Jacobi, M.T. Barbour, and E.W. Leppo. 2006. Benthic macroinvertebrate stream condition indices for New Mexico wadeable streams.

Jessup, B. 2009a. Development of Bedded Sediment Benchmarks for Oregon Streams. Prepared for Oregon Department of Environmental Quality and U.S. EPA Region 10.

Jessup B. 2009b. Recalibration of the Macroinvertebrate Multi-Metric Index for Colorado. Prepared for the Colorado Department of Public Health and Environment, Water Quality Control Division – Monitoring Unit, Denver, CO and EPA Region 8.

Kaufmann, P.R., D.V. Peck, J.M. Faustini, and S.G. Paulsen (in prep). Quantifying Habitat Quality for National Assessments.

Kaufmann, P.R., J.M. Faustini, D.P. Larsen, M. Shirazi. 2008. A Roughness-Corrected Index of Relative Bed Stability for Regional Stream Surveys. Geomorphology 99:150-170.

Kaufmann, P.R., D.P. Larsen, and J.M. Faustini. 2009. Bed Stability and Sedimentation Associated with Human Disturbances in Pacific Northwest Streams. Journal of the American Water Resources Association (JAWRA) 45(2):434-459.

King, R. S., and C. J. Richardson. 2003. Integrating bioassessment and ecological risk assessment: An approach to developing numerical water-quality criteria. Environmental Management 31:795-809.

Lazorchak, J.M., D.J. Klemm and D.V. Peck (eds.). 1998. Environmental Monitoring and Assessment Program -- Surface Waters: Field Operations and Methods for Measuring the Ecological Condition of Wadeable Streams. EPA/620/R-94/004F. U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C. 211pp plus Appendices.

Lisle, T. 1989. Sediment Transport and Resulting Deposition in Spawning Gravels, North Coast California. Wat. Resourc. Res. 25 (6):1303-1319.

Newcombe, C.P., and J.O.T. Jensen. 1996. Channel suspended sediment and fisheries: a synthesis for quantitative assessment of risk and impact. North American Journal of Fisheries Management. 16:693-727.

NMED (New Mexico Environment Department). 2009. Procedures for assessing water quality standards attainment for the state of New Mexico CWA §303(d) /§305(b); Integrated report: Assessment protocol. Appendix D: Sedimentation/Siltation Assessment Protocol for Wadeable, Perennial Streams. Surface Water Quality Bureau. Public comment draft April 20, 2009.

NMWQCC (New Mexico Water Quality Control Commission). 2007. Standards for Interstate and Intrastate Surface Waters. New Mexico Administrative Code.

Opperman, J.J., K.A. Lohse, C. Brooks, N.M. Kelly, and A.M. Merenlender. 2005. Influence of land use on fine sediment in salmonid spawning gravels within the Russian River Basin, California. Canadian Journal of Fisheries and Aquatic Sciences 62: 2740–2751.

Paul, J.F., and M.E. McDonald. 2005. Development of empirical, geographically specific water quality criteria: A conditional probability analysis approach. Journal of the American Water Resources Association (JAWRA) 41(5):1211-1223.

Paul, J. F., S. M. Cormier, W. J. Berry, P. R. Kaufmann, R. L. Spehar, D. J. Norton, R. E. Cantilli, R. Stevens, W. F. Swietlik, and B. K. Jessup. 2008. Developing Water Quality Criteria for Suspended and Bedded Sediments—Illustrative Example Application. Water Practice 2(1):1-17.

Paul, M.J. 2008. Draft Predictive Bioassessment Models for New Mexico Streams. Prepared for New Mexico Environment Department, Santa Fe, NM. Prepared by Tetra Tech, Inc., Owings Mills, MD.

Peck, D.V., A.T. Herlihy, B.H. Hill, R.M. Hughes, P.R. Kaufmann, D.J. Klemm, J.M. Lazorchak, F.H. McCormick, S.A. Peterson, P.L. Ringold, T. Magee, and M. Cappaert, 2006. Environmental Monitoring and Assessment Program-Surface Waters Western Pilot Study: Field Operations Manual for Wadeable Streams. EPA ⁄ 620 ⁄R-06 ⁄ 003. U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C. Available at: http://www.epa.gov/wed/pages/publications/authored/EPA620R-06003EMAPSWFieldOperationsManualPeck.pdf

Qian, S.S., King, R.S., Richardson, C.J., 2003. Two statistical methods for the detection of environmental thresholds. Ecological Modeling 166: 87-97.

R Development Core Team. 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Available at: http://www.R-project.org

Shuryn, D. 2007. Monitoring and Assessment of Sedimentation in Stream Channels of New Mexico. Master's Thesis, Water Resources Program, University of New Mexico.

Stoddard, J. L., D. V. Peck, A. R. Olsen, D. P. Larsen, J. Van Sickle, C. P. Hawkins, R. M. Hughes, T. R. Whittier, G. Lomnicky, A. T. Herlihy, P. R. Kaufmann, S. A. Peterson, P. L. Ringold, S. G. Paulsen, and R. Blair. 2005. Environmental Monitoring and Assessment Program (EMAP): western streams and rivers statistical summary. EPA 620/R-05/006. Office of Research and Development, US Environmental Protection Agency, Washington, DC.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson, and R. H. Norris. 2006. Setting expectations for the ecological condition of streams: the concept of reference conditions. Ecological Applications 16:1267–1276.

Suttle, K.B., M.E. Power, J.M. Levine, C. McNeely. 2004. How fine sediment in riverbeds impairs growth and survival of juvenile salmonids. Ecological Applications 14(4): 969–974.

U.S. EPA. 2000. Nutrient Criteria Technical Guidance Manual; Rivers and Streams. Office of Water, office of Science and Technology. EPA-822-B-00-002.

U.S. EPA. 2006. Framework for developing suspended and bedded sediment (SABS) water quality criteria. Office of Water, Office of Research and Development. EPA-822-R-06-001.

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

Society. 251 p.

Wood, P.J. and P.D. Armitage 1997. Biological effects of fine sediment in the lotic environment. Environmental Management 21:203–217.

Zar, J.H. 1999. Biostatistical Analysis. 4th edition. Prentice Hall. Upper Saddle River, New Jersey. 663pp. + appendices and literature cited.

Appendix A – GIS Procedures

Description of GIS Watershed Delineations and Accompanying Landscape Computations

The GIS work was completed by EPA Region 6 staff, Mr. Angel Kosfiszer [(214) 665-2187] and Robert Kirkland.

Data and watershed delineations were from (1) NM Rivpacs by TetraTech (Chris Warton), (2) NMED/Tetratech (Shann Stringer), (3) NMED (Seva Joseph), (4) WSA-SWIMS by Corvalis (Marlys Cappaert), and (5) CO by CO DPHE (Chris Theel).

A quick check of the 238 sampling sites, a number of watersheds showed errors in the delineations. Due to the difficulty of checking the accuracy of each 238 delineated watershed, all 238 watersheds were delineated.

After mapping each of the 238 sampling sites, it was observed that some of the sampling sites were less than 1 kilometer apart. In some instances, the close proximity of these sampling sites were acceptable since there were circumstances where there was a tributary between the sampling sites on the same main stream, or one sampling site was on a tributary and the other sampling site was on the main stream. However, for those paired sampling sites within <1km proximity of each other, as identified in **Table A-1**, each paired site data was reviewed and the sampling sites with the greater amount of data was selected in lieu of the other paired sampling site.

Based on the information described, 6 sampling sites were dropped leaving 232 watersheds to be both delineated and accompanying landscape computations as outlined in **Table A-2.**

The 232 watersheds were delineated using NHD+ Basins Delineation Tool. For those watersheds that were *nested*, many of the computations of the landscape parameters required individual watershed handling. Any parameters computed using the actual sampling site (point) was computed for all of the sampling sites at once. Since most GIS tools do not work well with nested watersheds, there was no good solutions to handle 232 watersheds landscape parameters in one single process

Table A-1. Paired sampling sites <1km apart

Decision	Distance (m)	Pairs	Station ID	Waterbody Name	Latitude	Longitude	Ecoregion	State
254.	3	01A	04ChicoC010.9	Chicorica Creek	36.77015	-104.396	26	NM
		01B	04UnaGat000.1	Una de Gato Creek	36.77243	-104.396	26	NM
142.	1	02A	28Comanc000.1	Comanche Creek	36.83114	-105.318	21	NM
		02B	28RCosti032.5	Rio Costilla 36.832	4	-105.318	21	NM
1.	7	03A	29RTusas000.1	Rio Tusas	36.38362	-106.036	21	NM
Deleted	03B		29RTusas000.2	Rio Tusas	36.3836	-106.036	21	NM
	688.0	04A	77EFkGil010.0	East Fork Gila River	33.18412	-108.165	23	NM
		04B	77EFkGil012.1	East Fork Gila River	33.18601	-108.158	23	NM
282.	3	05A	78GilaRi025.5	Gila River 32.	64927	-108.847	24	NM
Deleted	05B		78GilaRi026.1	Gila River 32.	64687	-108.846	24	NM
Deleted	268.1 06A		AZ06631-077	Mineral Creek	34.18167	-109.621	23	AZ
		06B	WAZP04-RMIN1	Mineral Creek	34.18004	-109.618	23	AZ
774.9		07A	AZ06631-125	Little Colorado River 34.007	78	-109.454	23	AZ
		07B	AZ06631-157	East Fork Little Colorado 34.	00139	-109.457	23	AZ
587.	8	08A	EPA01-0214	Rud1	34.010565	-109.28	23	AZ
Deleted	08B		AZ06631-237	Rudd Creek	34.00972	-109.274	23	AZ
	669.2	09A	AZ06631-110	Show Low Creek	34.17167	-109.983	23	AZ
		09B	AZ06631-186	Show Low Creek	34.17722	-109.986	23	AZ
300.	8	10A	CO06RS		37.81833	-107.721	21	CO
		10B	CO06TS		37.82056	-107.719	21	CO
Deleted	769.2 11A		CO03TS		37.85583	-107.576	21	CO
		11B	CO134M		37.86222	-107.573	21	CO
	553.0 12A		WCOP03-R007	East Fork Hermosa Creek 37.	63183	-107.879	21	CO
Deleted	12B		CO129M		37.63167	-107.873	21	CO
	482.9	13A	WAZP99-0590	Bright Angel Canyon	36.11783	-112.083	22	AZ
		13B	WAZP99-0768	Bright Angel Creek	36.1214	-112.08	22	AZ

Table A-2. Landscape variables and the scale of GIS analysis required.

Variable	Point adequate	Circle at least	Watershed required
Watershed area			✓
Stream order	✓		
Lithology		✓	
Precipitation	✓		
Elevation	✓		
Stream slope	✓		
Land slope		✓	
Ecoregion	✓		
Soils		✓	
Road density		✓	
Land use and cover		✓	
Dams and Diversions		✓	

The following was the sequential process for delineating the landscape parameters:

1. Downloaded all required data for the landscape computations, including NHD+ data for:10c ,11b,11d,13a,13b,14a,15a,15b
2. Computed Area (on the watershed) with XTools extension on ARCGIS.
3. Computed Stream Order (at the site) spatially joining sites with NHD+ files.
4. Computed Lithology (at the site) using State-specific USGS data.
5. Computed Precipitation (on the watershed) by clipping PRISM data (polygons) with Basin Delineations and calculating averages for Basin Area.
6. Computed Elevation (at the site) using NHD+ DEM files.
7. Computed Stream Slope (at the site) using NHD+ files and spatially joining sites with slope computations from NHD+ DEM files.
8. Computed Land Slope (on the watershed) using NHD+ and DEM. The values are percent grade.
9. Computed Ecoregion level 4 and 3(when 4 not available) (at the site) spatially joining EPA Ecoregion level 4 and 3 with site data.
10. Computed Soils Permeability (on the watershed) by clipping STATSGO data (polygons) with site Basin Delineations and calculating averages for Basin Area.
11. Computed Road Density (on the watershed) using Attila tool and TIGER 2000 files and the number of road/stream crossings using the tools in ARCGIS. First attempt to compute the number of crossings of roads and streams in the watershed showed that a lot of processing time was required by the Atilla tool. Because the process is one watershed at a time it was dimmed too costly. Found a more efficient method using ARCGIS and computed the road/stream crossings.
12. Computed Land Use and Cover (on the watershed) using Attila tool and NLCD 2001 data. The computed parameters (see Metadata for Landscape Characteristics that follows) are comprehensive enough that very little would be gained by computing population based on Census (e.g Nindex, Uindex, Purb).
13. Computed Dams (on the watershed) clipping US Army Corps of Engineers dam location data with Basin Delineations.

Metadata for Landscape Characteristics

Land_area - Total terrestrial area in map units (total area minus water)
LC_overlap - Percent overlap between reporting unit and land cover themes
SL_LndArea - Total terrestrial area (total area minus water) in map units for the land cover/slope composite grid
SL_Overlap - Percent overlap between reporting unit and land cover/slope composite grid

Land cover proportions
Pagc - Percentage of reporting unit that is crop land
Pagp - Percentage of reporting unit that is pasture
Pagt - Percentage of reporting unit that is all agricultural use
Pfor - Percentage of reporting unit that is forest
Pmbar - Percentage of reporting unit that is man made barren
Pnbar - Percentage of reporting unit that is natural barren
Png - Percentage of reporting unit that is natural grassland
Pshrb - Percentage of reporting unit that is shrubland
Purb - Percentage of reporting unit that is urban
Pusr - Percentage of reporting unit that is user defined class
Pwetl - Percentage of reporting unit that is wetland
N_index - Percentage of reporting unit that is all natural land use
U_index - Percentage of reporting unit that is all human land use

Each of the above will also have a field with _A appended (e.g. Pfor_A) representing total area in map units (meters).

Appendix B
Lithologic Erodibility

Lithologic units in and around New Mexico were determined using state-specific GIS data sets. The units were narrative descriptions of the lithologic rack types at five levels of detail (e.g., Sedimentary-Clastic-Sandstone-Arenite-Calcarenite). Similar lithologic types were found in the USGS Open-File Report 2005-1351 (accessible at: http://pubs.usgs.gov/of/2005/1351/). Each lithologic type was assigned attributes based on average values reported in the literature. This analysis was performed by John Olsen of the Western Center for Monitoring and Assessment of Freshwater Ecosystems. The variable related to lithologic erodibility was Average Uniaxial Compressive Strength (WtAvgUCS) in megapascals (MPa).

A review of rock hardness and type allowed expert erodibility rating on a scale of 1 (highly resistant to degradation) to 10 (likely to generate fine sediments during weathering or disturbance) (**Table B-1**). The experts who assigned the ratings included Ben Jessup and Nick Jokay of Tetra Tech and James Hogan of NMED. It was decided that ratings of 6 and higher indicated generally "erodible" rock types.

The erodibility ratings were used in GIS analyses of the percent of erodible rock types in the catchments of the sampling sites. The percentage of rock types with erodibility ratings of 6 or greater were calculated in each delineated catchment. "Percent erodible lithology" was used to estimate the effect of lithology on bedded stream sediment characteristics.

Table B-1. Erodibility ratings for rock types in and around New Mexico

Narrative Rank	ROCKTYPES	Erodibility Rating
Highly Resistant	alkaline basalt - basalt	3
Highly Resistant	andesite -	3
Highly Resistant	andesite - basalt	3
Highly Resistant	andesite - dacite	3
Highly Resistant	andesite - intermediate volcanic rock	3
Highly Resistant	basalt -	3
Highly Resistant	basalt - alkaline basalt	3
Highly Resistant	basalt - andesite	3
Highly Resistant	basalt - mafic volcanic rock	3
Highly Resistant	basalt - pyroclastic	3
Highly Resistant	basalt - rhyolite	3
Highly Resistant	dacite - rhyolite	3
Highly Resistant	diabase -	3
Highly Resistant	felsic gneiss - mafic gneiss	3
Highly Resistant	felsic metavolcanic rock - plutonic rock (phaneritic)	3
Highly Resistant	felsic volcanic rock - alkalic volcanic rock	3
Highly Resistant	felsic volcanic rock - intermediate volcanic rock	3
Highly Resistant	felsic volcanic rock - pyroclastic	3
Highly Resistant	gabbro - diorite	3

Table B-1. Erodibility ratings for rock types in and around New Mexico

Narrative Rank	ROCKTYPES	Erodibility Rating
Highly Resistant	granite - granodiorite	3
Highly Resistant	granitoid -	3
Highly Resistant	granitoid - diabase	3
Highly Resistant	greenstone - intermediate metavolcanic rock	3
Highly Resistant	mafic metavolcanic rock -	3
Highly Resistant	plutonic rock (phaneritic) -	3
Highly Resistant	rhyolite -	3
Highly Resistant	volcanic rock (aphanitic) -	3
Highly Resistant	volcanic rock (aphanitic) - intermediate volcanic rock	3
Highly Resistant	volcanic rock (aphanitic) - mixed clastic/volcanic	3
Resistant	biotite gneiss - schist	4
Resistant	granodiorite - granite	4
Resistant	granodiorite - granitoid	4
Resistant limestone	-	4
Resistant	limestone - dolostone (dolomite)	4
Resistant	limestone - fine-grained mixed clastic	4
Resistant	limestone - medium-grained mixed clastic	4
Resistant	limestone - quartzite	4
Resistant	limestone - sandstone	4
Resistant	limestone - shale	4
Resistant	limestone - siltstone	4
Resistant	metasedimentary rock -	4
Resistant	quartz latite -	4
Resistant	quartz monzonite - granitoid	4
Resistant	quartzite - slate	4
Resistant	rhyolite - dacite	4
Resistant	rhyolite - felsic volcanic rock	4
Resistant	andesite - tuff	5
Resistant	basalt - tuff	5
Resistant clastic	-	5
Resistant	clastic - carbonate	5
Resistant	clastic - mixed clastic/volcanic	5
Resistant	clastic - sedimentary rock	5
Resistant	clastic - volcanic rock (aphanitic)	5
Resistant mediu	m-grained mixed clastic - carbonate	5
Resistant	medium-grained mixed clastic - volcanic rock (aphanitic)	5
Resistant	mixed clastic/carbonate - evaporite	5
Resistant	mixed clastic/carbonate - limestone	5
Resistant	mixed clastic/volcanic - basalt	5
Resistant	phyllite - schist	5
Resistant	pyroclastic - tuff	5

Table B-1. Erodibility ratings for rock types in and around New Mexico

Narrative Rank	ROCKTYPES	Erodibility Rating
Resistant	rhyolite - tuff	5
Resistant	schist - gneiss	5
Erodible carbonate	-	6
Erodible	carbonate - clastic	6
Erodible	carbonate - fine-grained mixed clastic	6
Erodible	carbonate - medium-grained mixed clastic	6
Erodible	carbonate - sandstone	6
Erodible	conglomerate - sandstone	6
Erodible	fine-grained mixed clastic - carbonate	6
Erodible	fine-grained mixed clastic - coal	6
Erodible	fine-grained mixed clastic - dolostone (dolomite)	6
Erodible	fine-grained mixed clastic - evaporite	6
Erodible	fine-grained mixed clastic - limestone	6
Erodible	fine-grained mixed clastic - medium-grained mixed clastic	6
Erodible	fine-grained mixed clastic - sandstone	6
Erodible	fine-grained mixed clastic - sedimentary rock	6
Erodible	medium-grained mixed clastic -	6
Erodible	medium-grained mixed clastic - fine-grained mixed clastic	6
Erodible	mudstone - sandstone	6
Erodible sandstone-		6
Erodible sandstone	-	6
Erodible	sandstone - arkose	6
Erodible	sandstone - carbonate	6
Erodible	sandstone - claystone	6
Erodible	sandstone - conglomerate	6
Erodible	sandstone - fine-grained mixed clastic	6
Erodible	sandstone - limestone	6
Erodible	sandstone - medium-grained mixed clastic	6
Erodible	sandstone - mudstone	6
Erodible	sandstone - shale	6
Erodible	sandstone - siltstone	6
Erodible	sedimentary rock -	6
Erodible	sedimentary rock - medium-grained mixed clastic	6
Erodible shale	-	6
Erodible	shale - bentonite	6
Erodible	shale - claystone	6
Erodible	shale - fine-grained mixed clastic	6
Erodible	shale - limestone	6
Erodible	shale - sandstone	6
Erodible shale-siltstone		6
Erodible	siltstone - dolostone (dolomite)	6

Table B-1. Erodibility ratings for rock types in and around New Mexico

Narrative Rank	ROCKTYPES	Erodibility Rating
Erodible	siltstone - sandstone	6
Erodible	siltstone - shale	6
Erodible	clastic - unconsolidated deposit	7
Erodible	lava flow - tuff	7
Erodible	medium-grained mixed clastic - tuff	7
Erodible	tuff - pyroclastic	7
Erodible	landslide - talus	8
Highly erodible	alluvium - unconsolidated deposit	9
Highly erodible	ash-flow tuff -	9
Highly erodible	coarse-grained mixed clastic - fine-grained mixed clastic	9
Highly erodible	coarse-grained mixed clastic - unconsolidated deposit	9
Highly erodible	eolian -	9
Highly erodible	glacial drift -	9
Highly erodible	gravel - alluvium	9
Highly erodible	gravel - sand	9
Highly erodible	landslide - colluvium	9
Highly erodible	alluvial fan -	10
Highly erodible	alluvium -	10
Highly erodible	alluvium - eolian	10
Highly erodible	clay or mud - silt	10
Highly erodible	dune sand - silt	10
Highly erodible	lake or marine deposit (non-glacial) - alluvium	10
Highly erodible	sand-	10
Highly erodible	sand - clay or mud	10
Highly erodible	sand - conglomerate	10
Highly erodible	sand - gravel	10
Highly erodible	sand - silt	10
Highly erodible	sand-siltstone	10
Highly erodible	silt	10
Highly erodible	silt - sand	10
Highly erodible	terrace - sand	10
Highly erodible	unconsolidated deposit - sand	10
Not applicable	evaporite -	
Not applicable	evaporite - evaporite	
Not applicable	evaporite - fine-grained mixed clastic	
Not applicable	evaporite - limestone	
Not applicable	evaporite - sandstone	
Not applicable	indeterminate -	
Not applicable	water -	

Appendix C.
Distributions of bedded sediment indicators.

Codes in graphs include sediment indicators and reference status, as follows:

Code	Description
PCT_SaFn	Percent Sand & Fines
PCT_FN	Percent Fines
LRBS_fin	Log Relative Bed Stability
LRBS_NOR	Log Relative Bed Stability excluding bedrock and hardpan
trPCT_SaFn	Percent Sand & Fines
trPCT_ FN	Percent Fines
Resid_pSaFn	Residual Percent Sand & Fines based on reference relationship with critical particles diameter
Resid_pFN	Residual Percent Fines based on reference relationship with critical particles diameter
Res2pSaFn	Residual Percent Sand & Fines based on Western EMAP formula
Res2pFines	Residual Percent Fines based on Western EMAP formula
Ref	Reference sites
Other	"Other" sites, neither reference nor stressed
Strs	Stressed sites

Statistics in these graphs are identical to those presented in **Table 10** of the report.

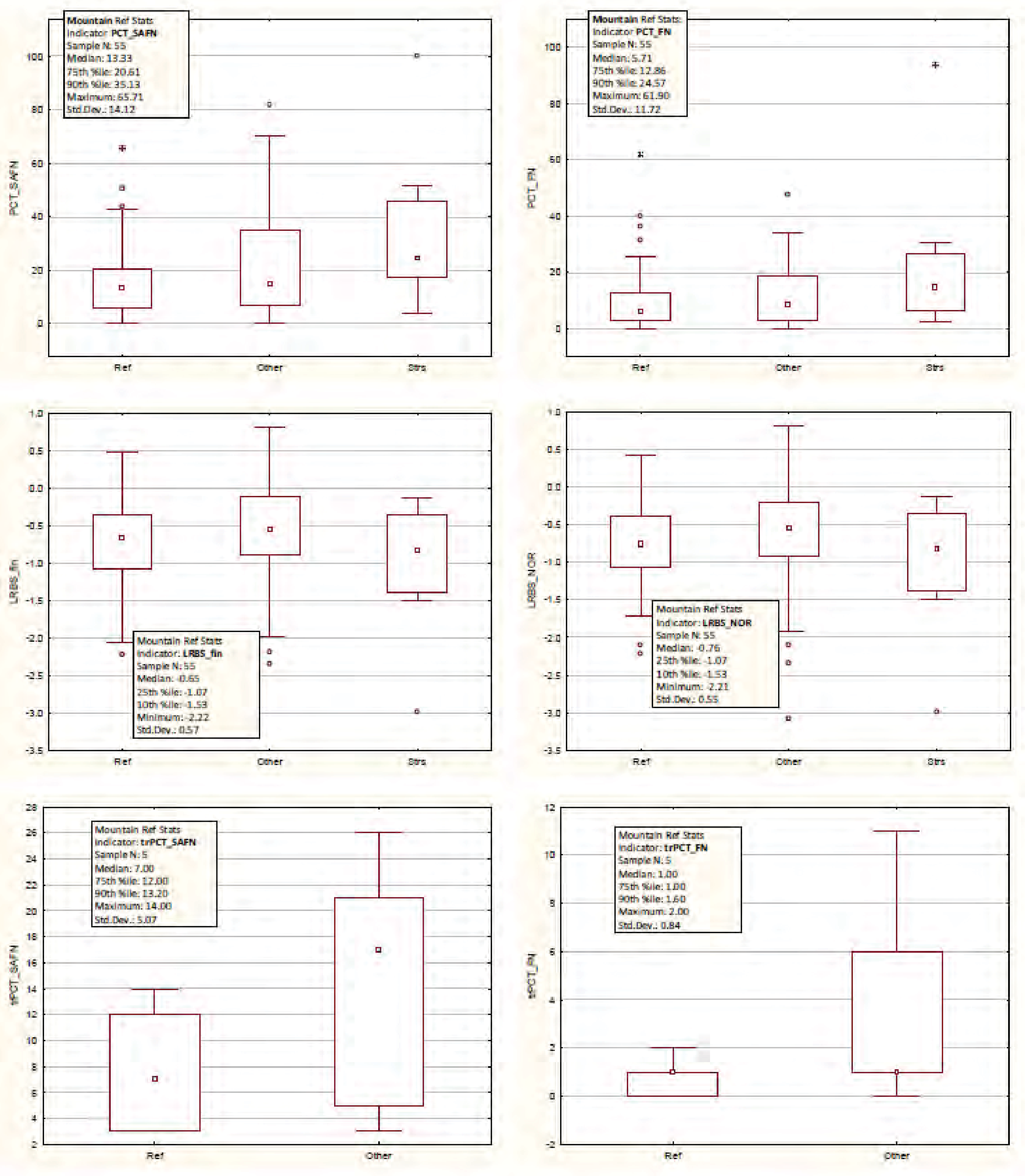

Figure C-1. Bedded sediment indicator distributions in Reference, Other, and Stressed sites in the Mountain site class, including statistics for Reference sites.

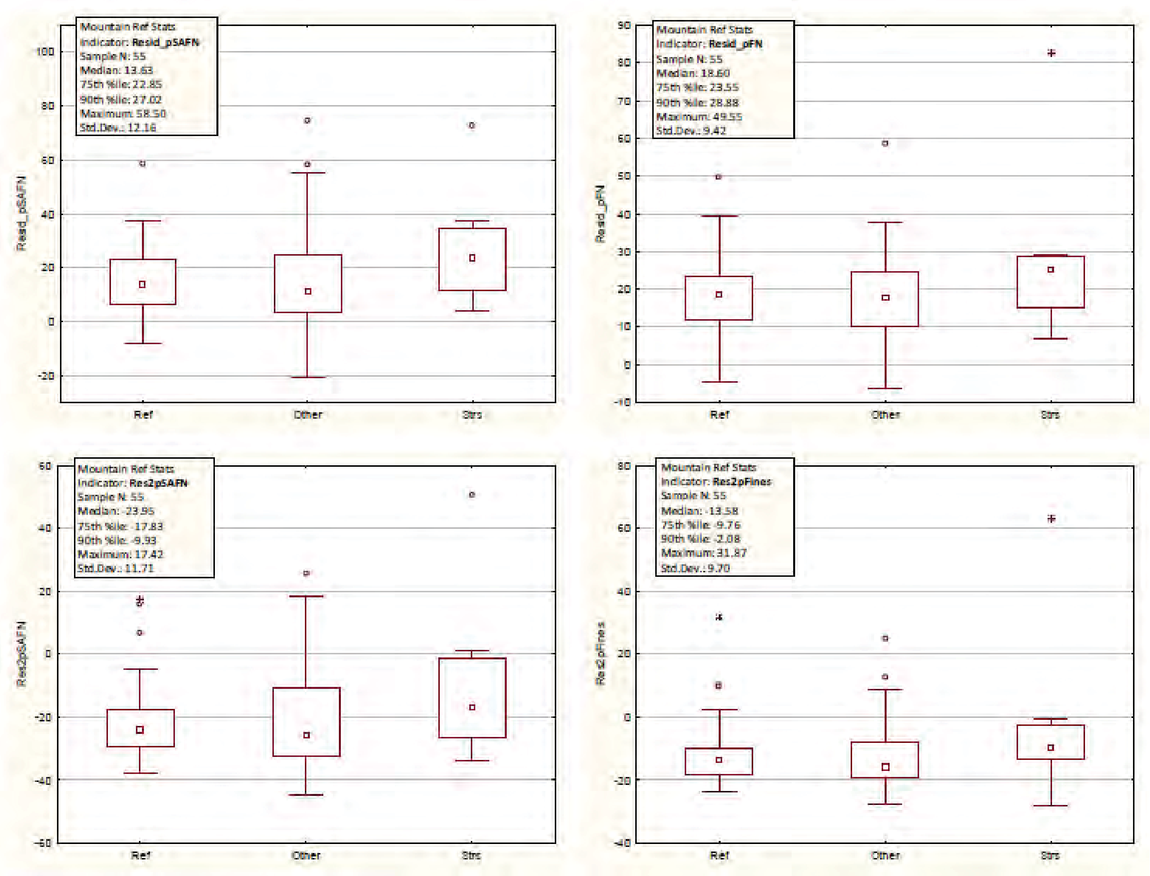

Figure C-1 (continued). Bedded sediment indicator distributions in Reference, Other, and Stressed sites in the Mountain site class, including statistics for Reference sites.

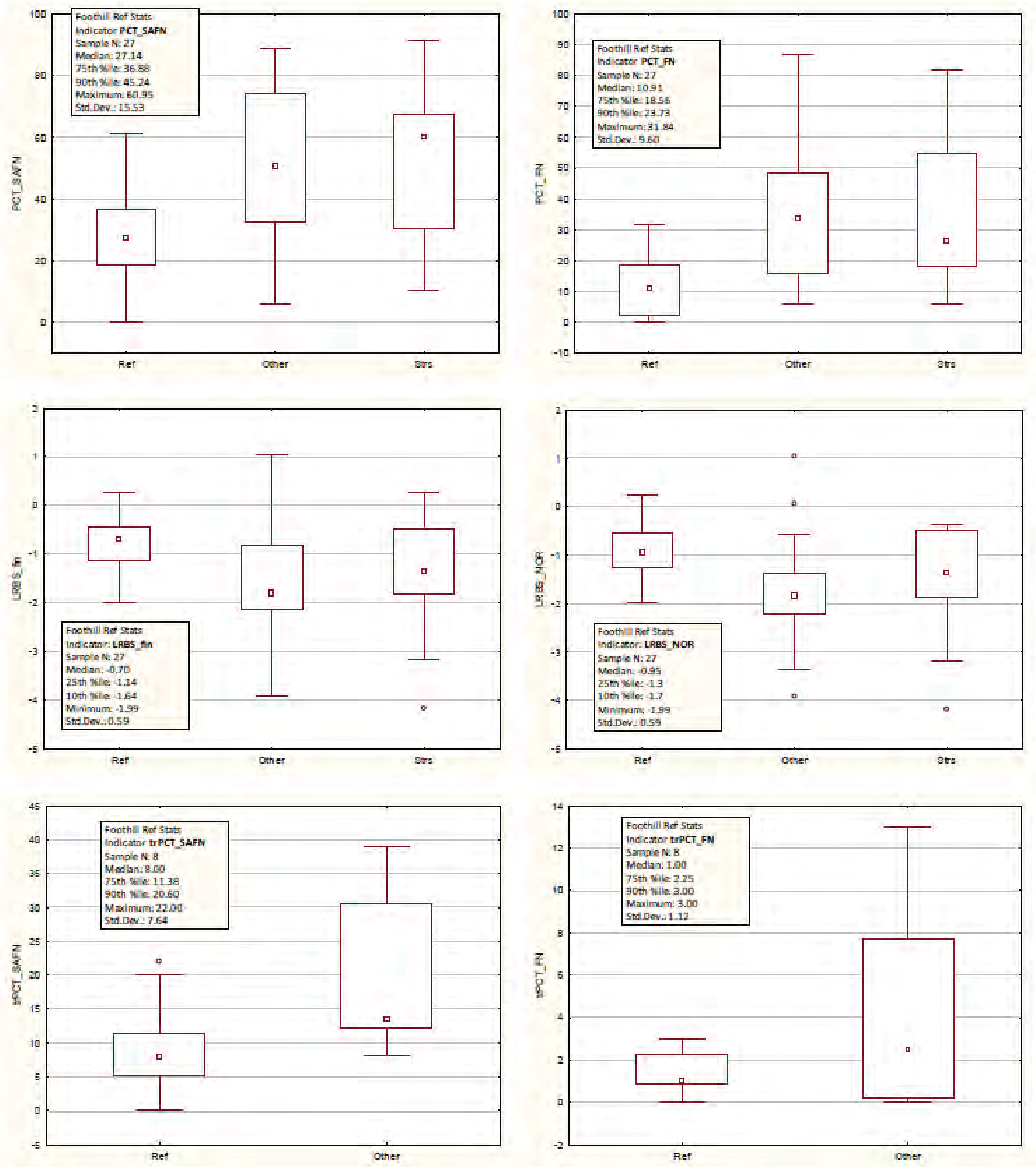

Figure C-2. Bedded sediment indicator distributions in Reference, Other, and Stressed sites in the Foothills site class, including statistics for Reference sites.

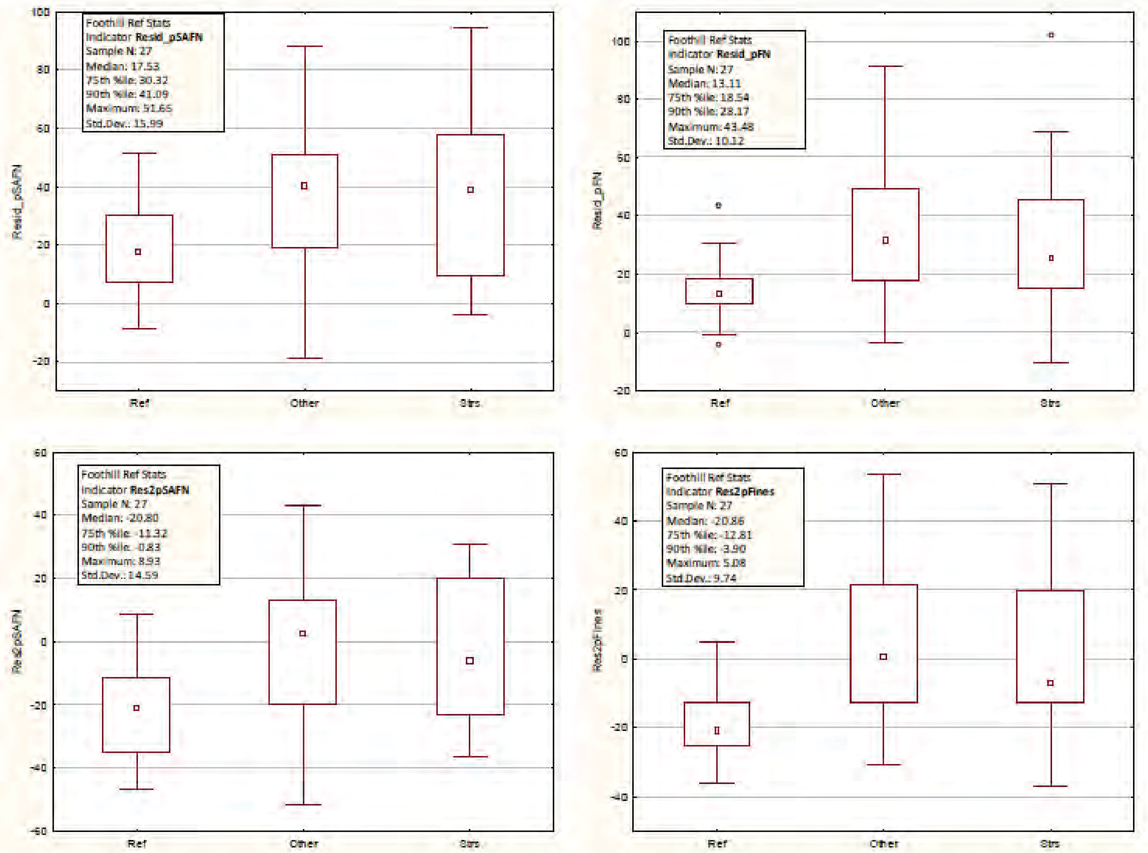

Figure C-2 (continued). Bedded sediment indicator distributions in Reference, Other, and Stressed sites in the Foothills site class, including statistics for Reference sites.

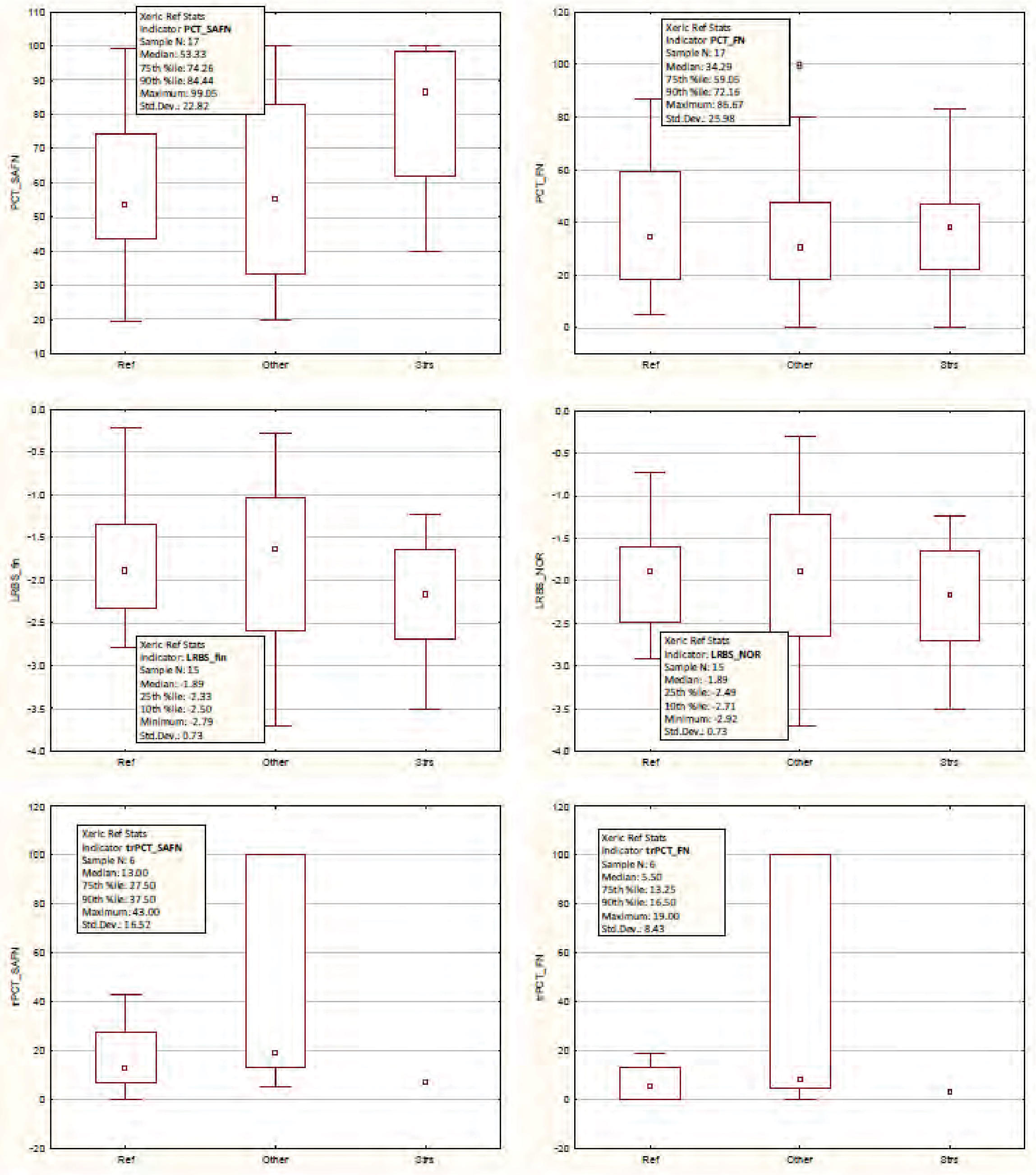

Figure C-3. Bedded sediment indicator distributions in Reference, Other, and Stressed sites in the Xeric site class, including statistics for Reference sites.

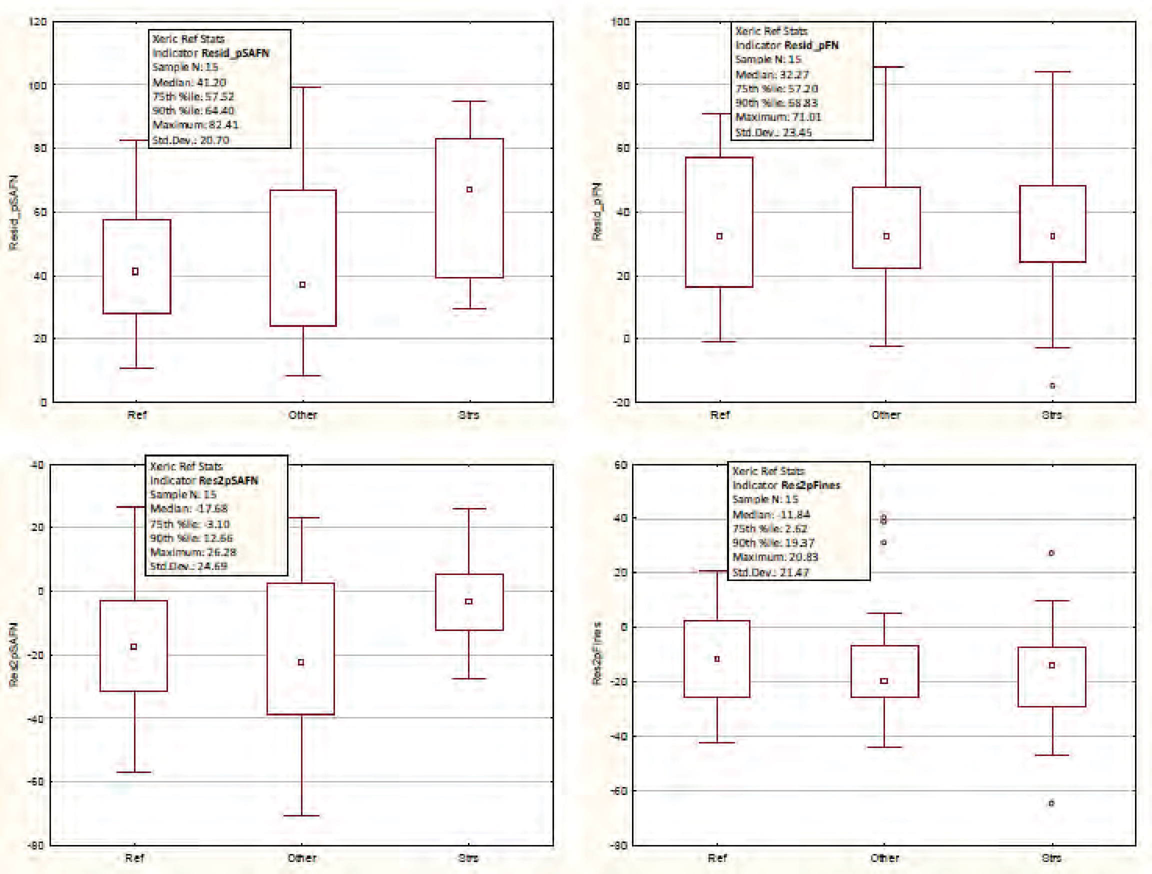

Figure C-3 (continued). Bedded sediment indicator distributions in Reference, Other, and Stressed sites in the Xeric site class, including statistics for Reference sites.

Codes in graphs include suspended and bedded sediment indicators, as follows:

Code Description	
LTurb Log	Turbidity
LTSS	Log Total Suspended Solids
SRO_Turb	Spring Run-off Turbidity (includes Storm Flow as solid markers)
SRO_TSS	Spring Run-off Total Suspended Solids (includes Storm Flow as solid markers)
PCT_SaFn	Percent Sand & Fines
PCT_FN Percent	Fines
LRBS_fin	Log Relative Bed Stability
Resid_pSaFn	Residual Percent Sand & Fines based on reference relationship with critical particles diameter
Resid_pFN	Residual Percent Fines based on reference relationship with critical particles diameter
Res2pSaFn	Residual Percent Sand & Fines based on Western EMAP formula
Res2pFines	Residual Percent Fines based on Western EMAP formula

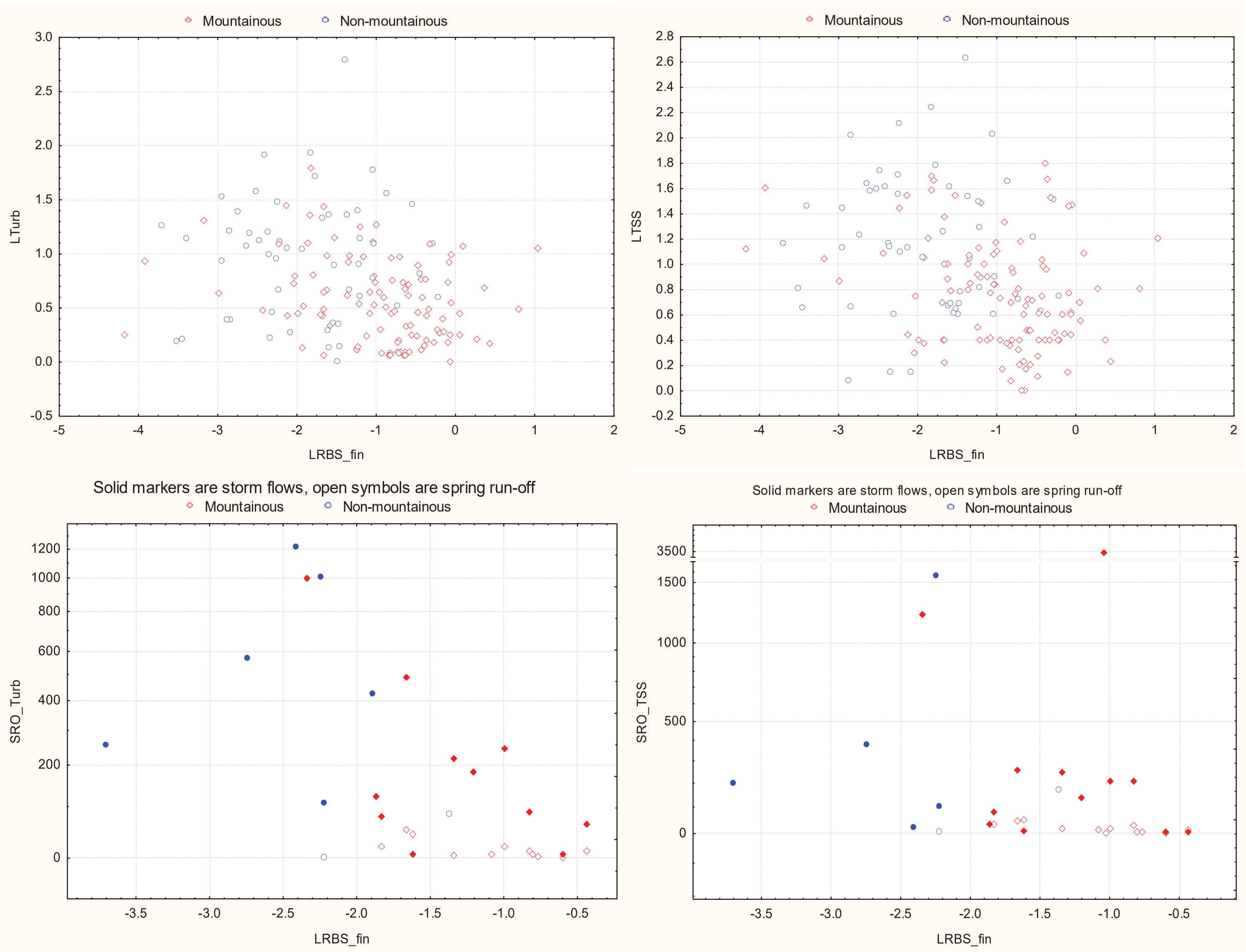

Figure D-1. Suspended sediment as a function of LRBS values. Mountains include ecoregions 21, 23 and 79. The top graphs are turbidity and TSS on a 'log10' scale during low flow events. Spring run-off and storm flows are in the lower graphs.

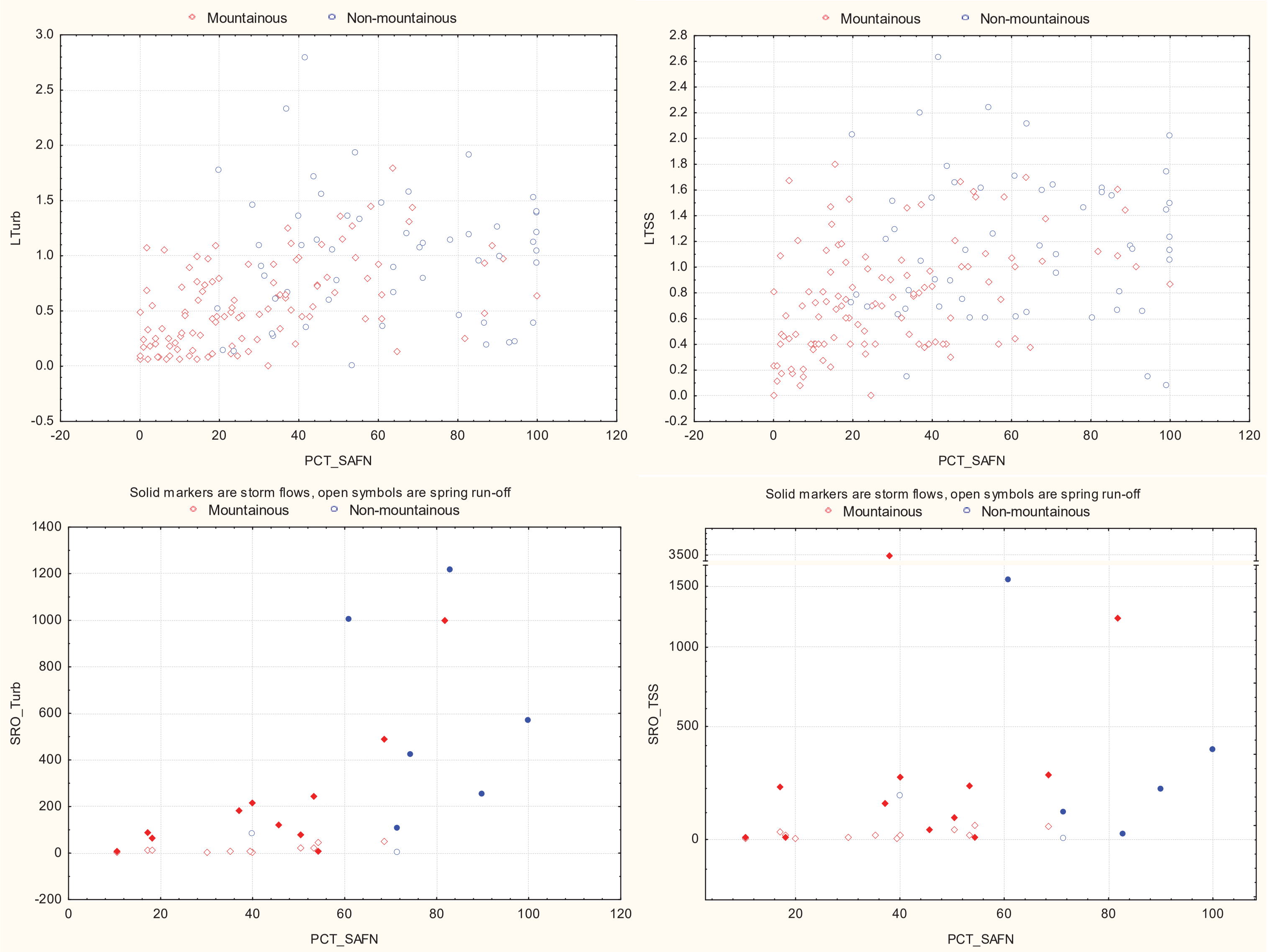

Figure D-2. Suspended sediment as a function of percent sand and fines. Mountains include ecoregions 21, 23 and 79. The top graphs are turbidity and TSS on a ‘log10’ scale during low flow events. Spring run-off and storm flows are in the lower graphs.

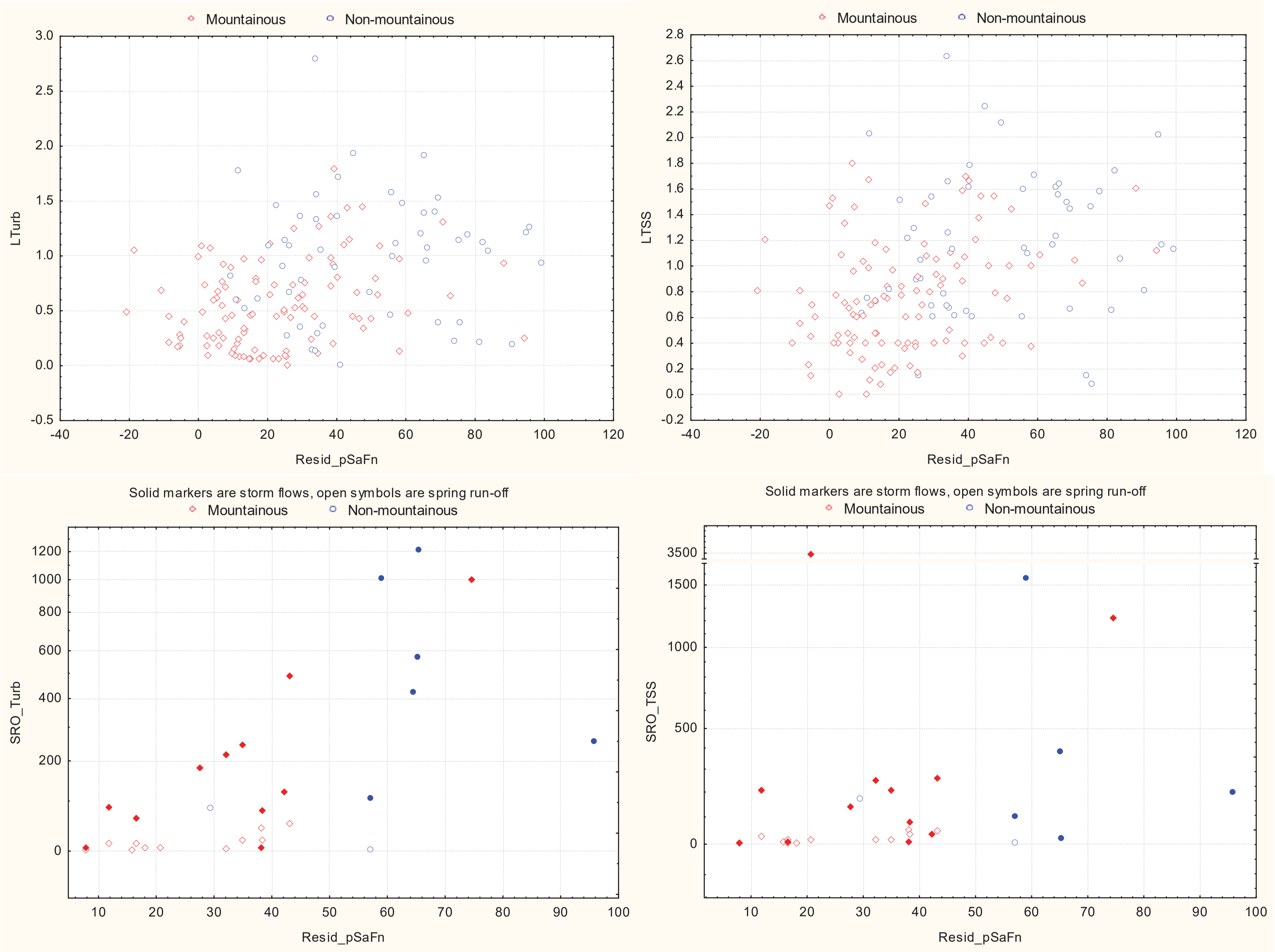

Figure D-3. Suspended sediment as a function of residual % sand and fines. Mountains include ecoregions 21, 23 and 79. The top graphs are turbidity and TSS on a ‘log10’ scale during low flow events. Spring run-off and storm flows are on the lower graphs.

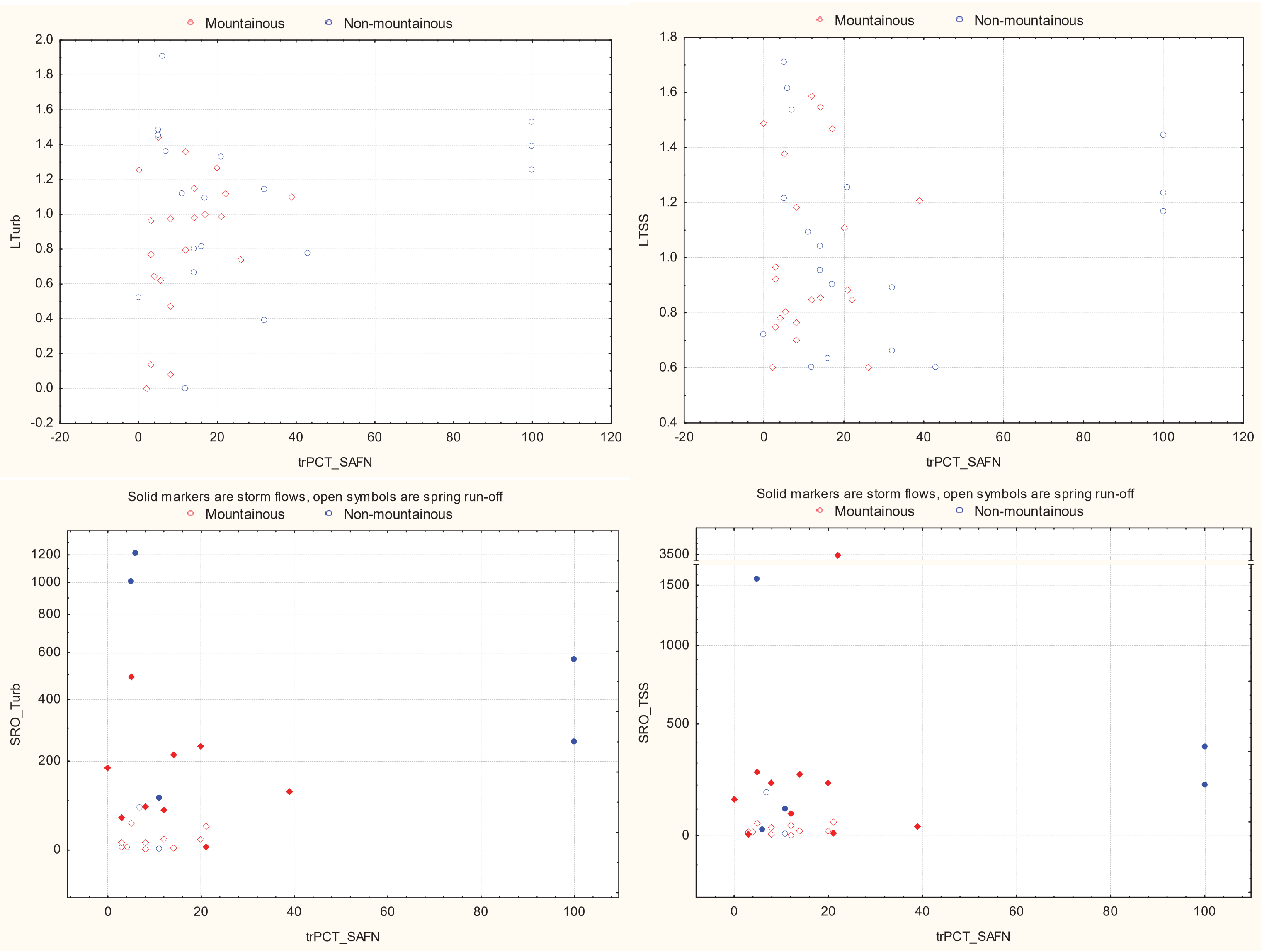

Figure D-4. Suspended sediment as a function of targeted riffle % sand and fines. Mountains include ecoregions 21, 23 and 79. The top graphs are turbidity and TSS on a ‘log10’ scale during low flow events. Spring run-off and storm flows are on the lower graphs.

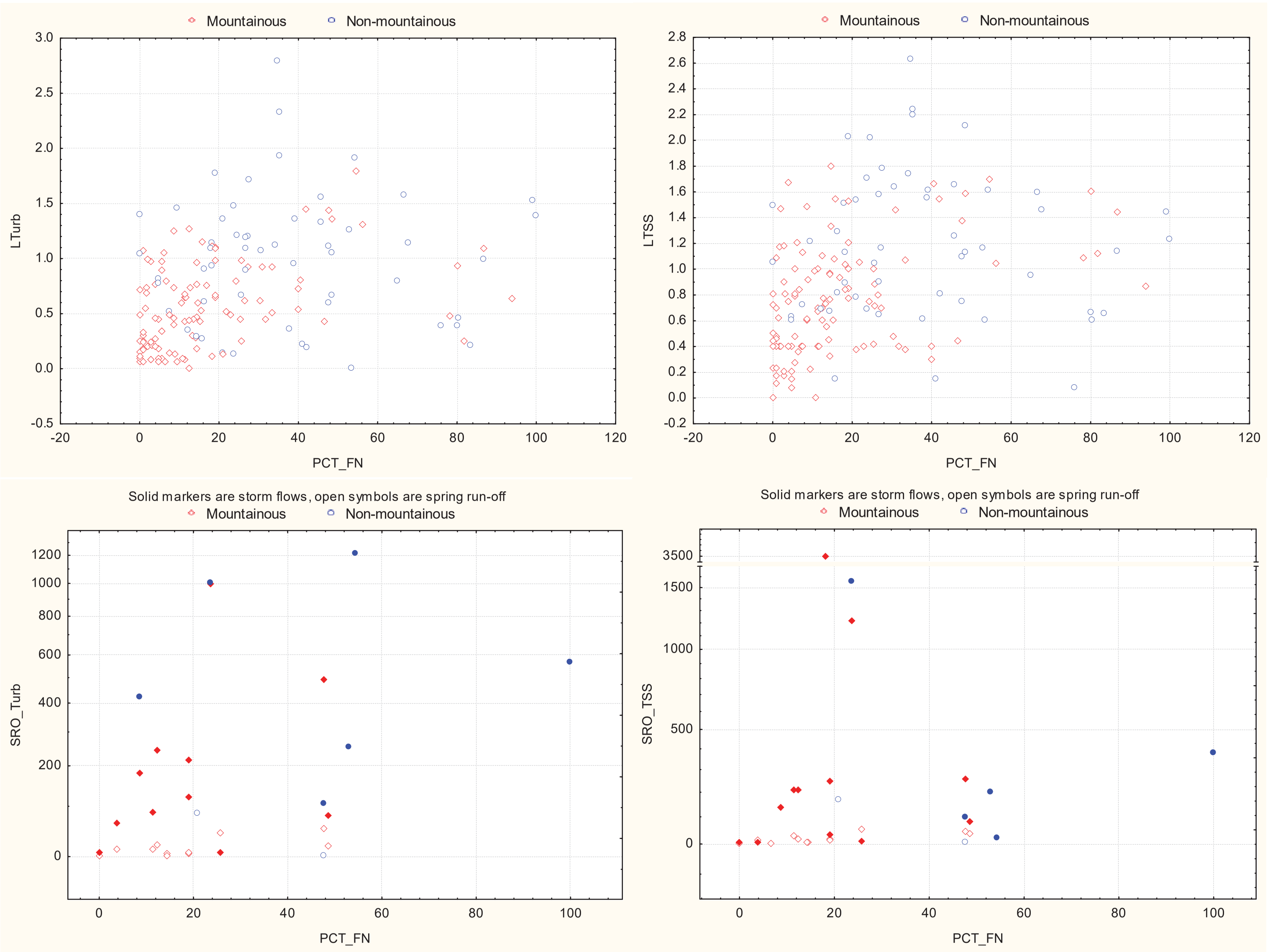

Figure D-5. Suspended sediment as a function of percent fines. Mountains include ecoregions 21, 23 and 79. The top graphs are turbidity and TSS on a ‘log10’ scale during low flow events. Spring run-off and storm flows are on the lower graphs.

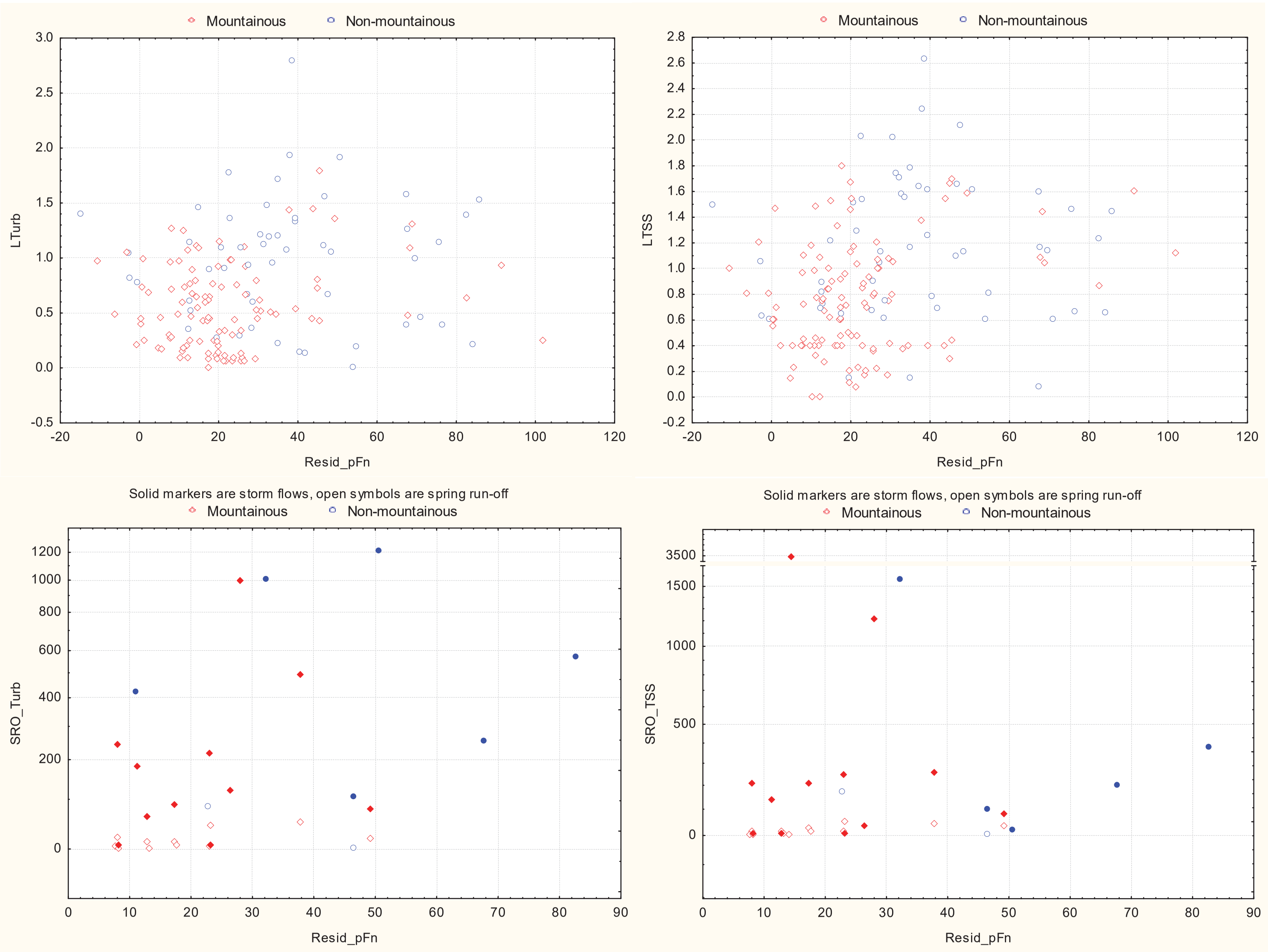

Figure D-6. Suspended sediment as a function of residual percent fines. Mountains include ecoregions 21, 23 and 79. The top graphs are turbidity and TSS on a ‘log10’ scale during low flow events. Spring run-off and storm flows are on the lower graphs.

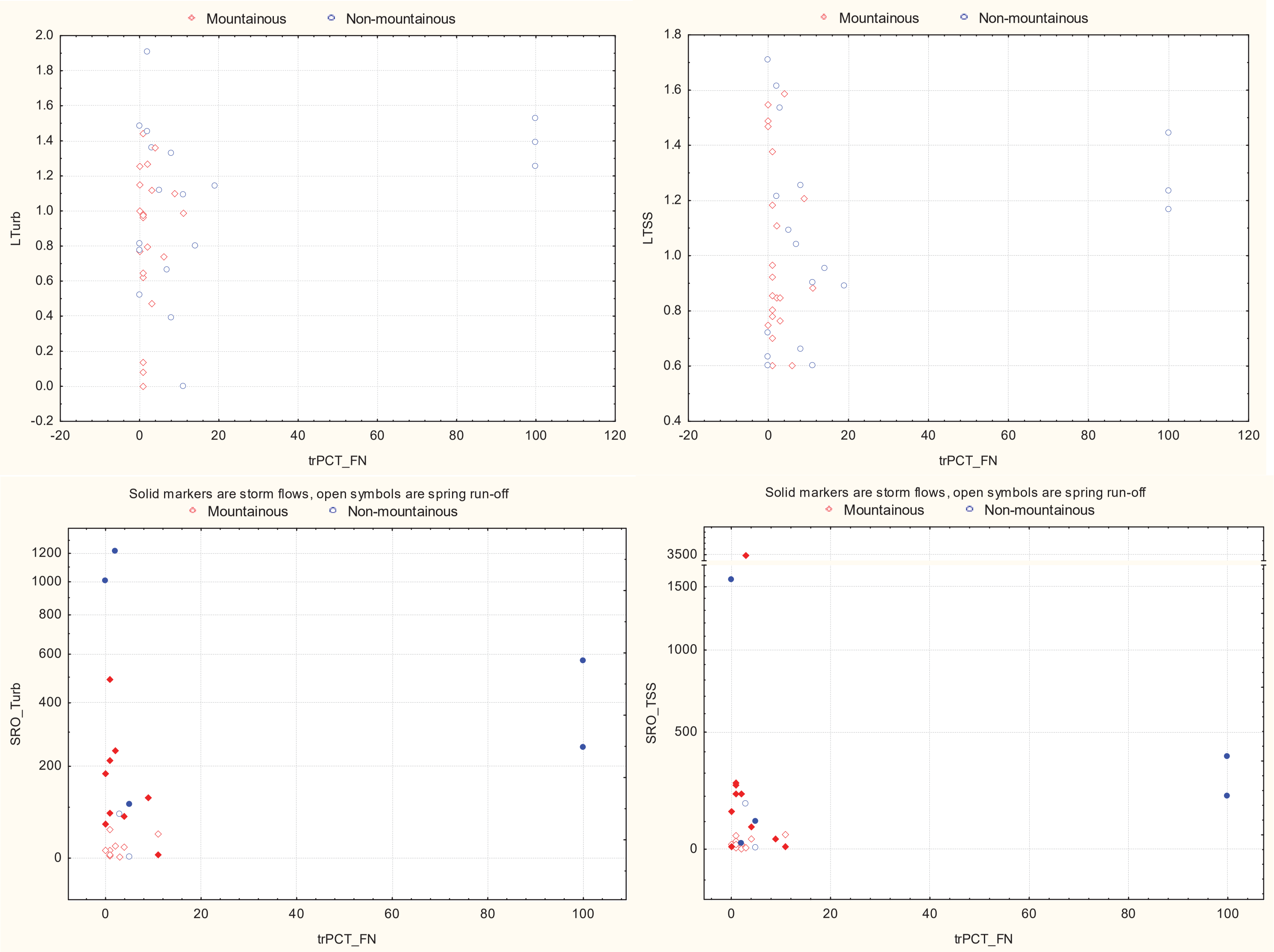

Figure D-7. Suspended sediment as a function of targeted riffle % fines. Mountains include ecoregions 21, 23 and 79. The top graphs are turbidity and TSS on a ‘log10’ scale during low flow events. Spring run-off and storm flows are on the lower graphs.

Table D-1. Correlations (Spearman's rho) among average suspended and bedded measures for streams in New Mexico. SD = standard deviation, SRO = spring run-off, strm = storm flow, tr = targeted riffle. Significant correlations (p<0.05) are bold-typed.

	TURB_ NTU	TSS_ MG_L	SD_ turb	SD_ TSS	SRO_ Turb	SRO TSS	strm_ Turb	strm_ TSS	LRBS fin	PCT FN	PCT SAFN	Tr PCT FN	Tr PCT SAFN
TURB_NTU	N=669	N=669	N=483	N=427	N=210	N=214	N=143	N=149	N=160	N=163	N=163	N=39	N=39
TSS_MG_L	**0.83**	N=680	N=483	N=430	N=210	N=214	N=144	N=150	N=163	N=166	N=166	N=39	N=39
SD_turb	**0.89**	**0.80**	N=483	N=424	N=204	N=208	N=133	N=139	N=53	N=54	N=54	N=35	N=35
SD_TSS	**0.74**	**0.93**	**0.76**	N=430	N=181	N=185	N=126	N=130	N=48	N=49	N=49	N=33	N=33
SRO_Turb	**0.49**	**0.47**	**0.30**	**0.47**	N=213	N=213	N=88	N=89	N=13	N=13	N=13	N=13	N=13
SRO_TSS	**0.43**	**0.47**	**0.26**	**0.45**	**0.90**	N=218	N=89	N=90	N=14	N=14	N=14	N=14	N=14
strm_Turb	**0.53**	**0.47**	**0.48**	**0.44**	**0.32**	**0.31**	N=152	N=152	N=16	N=16	N=16	N=14	N=14
strm_TSS	**0.47**	**0.46**	**0.40**	**0.43**	**0.34**	**0.37**	**0.90**	N=158	N=17	N=17	N=17	N=15	N=15
LRBS_fin	**-0.34**	**-0.36**	**-0.46**	-0.24	-0.31	-0.50	**-0.70**	-0.28	N=165	N=165	N=165	N=37	N=37
PCT_FN	**0.42**	**0.41**	**0.37**	0.28	0.31	**0.57**	**0.52**	0.18	**-0.72**	N=168	N=168	N=38	N=38
PCT_SAFN	**0.52**	**0.45**	**0.53**	0.26	0.33	0.46	**0.71**	0.31	**-0.84**	**0.76**	N=168	N=38	N=38
trPCT_FN	0.19	0.06	0.24	-0.01	0.09	0.19	-0.02	-0.10	**-0.51**	**0.70**	**0.58**	N=39	N=39
trPCT_SAFN	0.22	-0.01	0.17	0.03	0.10	0.21	0.01	0.17	-0.32	0.30	**0.50**	**0.62**	N=39

Appendix E
Assessment Results for Bedded Sediments

In **Table E-1**, assessment results refer to the two tiered assessment procedure suggested in **Figures 47-49** and **Table 16** of the report.

Table E-1. Assessment results for bedded sediment sites.

Site Code	SITENAME	RefStat	SiteClass	Tier 1 PCT_SAFN	Tier 2 LRBS_NOR	Assessment
02DryCim047.2	Dry Cimarron River	Ref	Xer	71.28		1P
02DryCim074.5	Dry Cimarron River	Other	Xer	99.05	-2.95	Fail
04ChicoC010.9	Chicorica Creek	Other	Xer	80.95	-2.61	Fail
04RatonC007.8	Raton Creek	Other	Xer	55.24	-1.75	Pass
04UnaGat000.1	Una de Gato Creek	Other	Xer	100.00	-2.74	Fail
05Cieneg006.3	Cieneguilla Creek	Other	Mtn	68.57	-1.66	Fail
05MPonil000.1	Middle Ponil Creek	Ref	Mtn	40.00	-1.34	Fail
05NPonil000.1	North Ponil Creek	Ref	FtHill	53.33	-1.00	1F, 2P
05PonilC000.1	Ponil Creek	Other	Xer	82.86	-2.48	1F, 2P
05PonilC014.9	Ponil Creek	Other	Xer	71.43	-2.22	Pass
05RAYAD038.4	Rayado Creek 3 miles above NM 21	Other	Xer	27.62	-0.94	Pass
05Rayado001.8	Rayado Creek	Other	Xer	90.00	-3.70	Fail
06Canadi305.0	Canadian River	Other	Xer	28.42	-0.71	Pass
10UteCre104.3	Ute Creek	Ref	Xer	44.76	-1.24	Pass
10UteCre150.7	Ute Creek	Ref	Xer	53.33	-2.52	1P, 2F
16Seneca043.0	Seneca Creek	Ref	Xer	80.41	-2.51	Fail
27RPinos007.3		Ref	Xer	19.52	-0.73	Pass
28Comanc000.1	Comanche Creek	Other	Mtn	39.42	-0.81	1F, 2P
28RCosti032.5	Rio Costilla	Ref	Mtn	20.00	-1.03	Pass
28RGRanc013.1	Rio Grande del Rancho	Ref	FtHill	36.62	-1.37	1P, 2F
28RSanBa013.2	Rio Santa Barbara	Ref	Mtn	10.48	-0.60	Pass
28SanCru004.2	Rio Santa Cruz	Strs	Xer	40.00	-1.37	Pass
28SanCru012.1	Rio Santa Cruz	Other	Xer	37.14	-1.34	Pass
29Abiqui002.3		Other	FtHill	56.60	-1.67	Fail
29Cecili000.1	Cecilia Creek	Other	Mtn	44.76	-0.66	1F, 2P
29Chihua001.3	Chihuahuenos Creek	Ref	Mtn	50.91	-2.22	Fail
29ClearC000.1	Clear Creek	Other	Mtn	54.29	-1.62	Fail
29Coyote017.5	Coyote Creek	Ref	Mtn	16.36	-1.01	Pass
29ElRito035.9	El Rito	Other	Mtn	1.82	0.37	Pass
29ElRito050.2	El Rito	Ref	Mtn	12.73	-1.12	1P, 2F
29LitTus003.4	Little Tusas	Other	Mtn	16.36	-0.09	Pass

Site Code	SITENAME	RefStat	SiteClass	Tier 1 PCT_SAFN	Tier 2 LRBS_NOR	Assessment
29Polvad009.0	Polvadera Creek	Other	FtHill	57.34	-2.03	Fail
29RCanji039.4	Rio Canjilon	Other	Mtn	15.61	-0.39	Pass
29RCHAMA143.8	Rio Chama	Other	FtHill	18.10	-0.78	Pass
29RChama183.4	Rio Chama	Other	FtHill	17.14	-0.83	Pass
29RChami002.7	Rio Chamita	Other	FtHill	50.48	-1.83	Fail
29RGalli005.5	Rio Gallina	Other	Mtn	81.82	-2.34	Fail
29RGalli048.3	Rio Gallina	Ref	Mtn	19.21	-0.81	Pass
29RioOso004.7	Rio del Oso	Ref	Xer	74.26	-1.89	1F, 2P
29RMedio002.7	Rio del Medio	Other	Mtn	27.27		1F, 2P
29RNutri027.5	Rio Nutrias	Strs	FtHill	63.64	-1.88	Fail
29RPuerc037.5	Rio Puerco de Chama	Ref	Mtn	18.10	-0.71	Pass
29RRESUM001.7	Rito Resumidero @ FR 93	Ref	Mtn	20.00	-0.69	Pass
29RTusas000.1	Rio Tusas	Other	FtHill	45.71	-1.86	Fail
29RTusas028.5	Rio Tusas	Ref	Mtn	43.64	-1.21	Fail
29RValle037.8	Rio Vallecito	Ref	Mtn	3.01	-0.09	Pass
40Alamos058.5	Alamosa Creek	Ref	FtHill	38.10	-1.23	1F, 2P
41Anima029.3	Las Animas Creek	Ref	FtHill	27.14	-0.70	Pass
41Percha025.3	Percha Creek	Other	Xer	31.43	-0.45	Pass
50PecosR670.3	Pecos River	Other	Xer	40.70	-1.19	Pass
60BlackR023.7	Black River	Other	Xer	30.00	-1.42	Pass
60BlackR052.0	Black River	Other	Xer	86.67	-3.03	Fail
60BlueSp002.0	Blue Spring	Ref	Xer	75.24	-2.83	Fail
60Sittin001.6	Sitting Bull Creek	Other	FtHill	32.38	-2.05	1P, 2F
77BlackC028.3	Black Canyon Creek	Ref	Mtn	10.22	-0.26	Pass
77Bobcat000.8	Bobcat Spring	Ref	FtHill	39.05	-1.52	Fail
77Bonner002.4	Bonner Creek	Other	Mtn	10.48	-0.28	Pass
77CubCre005.6	Cub Creek	Ref	Mtn	9.52	-0.59	Pass
77Diamon033.2	Main Diamond Creek	Ref	Mtn	15.15	-0.15	Pass
77EFkGil000.2	East Fork Gila River	Ref	FtHill	35.24	-1.18	Pass
77EFkGil010.0	East Fork Gila River	Ref	FtHill	49.05	-1.62	Fail
77EFkGil012.1	East Fork Gila River	Ref	FtHill	60.95	-1.69	Fail
77GilaRi092.0	Gila River	Ref	FtHill	36.54	-0.95	Pass
77IronCr009.7	Iron Creek	Ref	Mtn	27.18	-1.24	Fail
77MFkGil028.3	Middle Fork Gila River	Ref	FtHill	25.71	-0.57	Pass
77WFkGil010.0	West Fork Gila River	Ref	FtHill	30.11	-0.82	Pass
77WFkGil038.1	West Fork Gila River	Ref	Mtn	11.43	-0.34	Pass
77Willow000.6	Willow Creek	Ref	Mtn	15.87	-0.63	Pass
78BearCr027.0	Bear Creek	Ref	FtHill	17.14	-1.02	Pass

Site Code	SITENAME	RefStat	SiteClass	Tier 1 PCT_SAFN	Tier 2 LRBS_NOR	Assessment
78BlueCr000.9	Blue River	Ref	Xer	49.52	-1.08	Pass
78GilaRi025.5	Gila River	Ref	Xer	60.91	-2.25	Pass
78GilaRi052.6	Gila River	Strs	FtHill	67.62	-3.18	Fail
78GilaRi069.2	Gila River	Ref	FtHill	37.14	-1.21	1F, 2P
78GilaRi074.8	Gila River	Ref	FtHill	42.70	-1.99	Fail
78GILARI087.7	Gila River at NM 211 Bridge	Other	FtHill	24.76	-0.99	Pass
80SanFra028.6	San Francisco River	Other	FtHill	51.03	-1.53	Fail
80SanFra154.1	San Francisco River	Other	Mtn	14.29	-0.05	Pass
AZ06631-037	Hall Creek	Ref	Mtn	42.86	-1.18	Fail
AZ06631-038	Morrison Creek	Other	Mtn	70.48	-2.10	Fail
AZ06631-050	Mill Creek	Other	Mtn	51.43	-3.08	Fail
AZ06631-053	Hall Creek	Ref	Mtn	22.86	-1.31	Fail
AZ06631-061	Silver Creek	Other	FtHill	23.48	-0.56	Pass
AZ06631-065	South Fork Little Colorado River at Campg	Ref	Mtn	13.33	-0.97	Pass
AZ06631-093	Show Low Creek	Other	Mtn	31.43	-0.85	1F, 2P
AZ06631-097	Little Colorado River	Strs	FtHill	36.00	-1.38	1P, 2F
AZ06631-098	Riggs Creek	Other	FtHill	79.05	-3.37	Fail
AZ06631-109	Hall Creek	Other	Mtn	2.86	-0.64	Pass
AZ06631-110	Show Low Creek	Strs	FtHill	30.48	-0.49	Pass
AZ06631-125	Little Colorado River	Other	Mtn	19.05	-0.92	Pass
AZ06631-130	Rudd Creek	Other	FtHill	74.29	-2.21	Fail
AZ06631-133	East Fork Little Colorado	Ref	Mtn	50.48	-1.72	Fail
AZ06631-137	Coyote Creek at Richville	Other	FtHill	80.95	-1.80	Fail
AZ06631-141	Benton Creek	Ref	Mtn	0.98	-0.60	Pass
AZ06631-145	Little Colorado River	Other	FtHill	78.10	-3.04	Fail
AZ06631-149	Silver Creek	Other	FtHill	34.55	-1.40	1P, 2F
AZ06631-155	Little Colorado River	Other	Xer	94.76	-2.16	1F, 2P
AZ06631-157	East Fork Little Colorado	Other	Mtn	13.19	-0.53	Pass
AZ06631-186	Show Low Creek	Strs	FtHill	10.48	-0.35	Pass
AZ06631-210	South Fork Little Colorado	Ref	Mtn	11.43	-0.82	Pass
AZ06631-LCLVL001.32	Lee Valley Creek	Ref	Mtn	65.71	-2.10	Fail
AZ06631-LCRCR340.02	Little Colorado River downstream of Eagar	Other	FtHill	43.27	-1.37	Fail
CC0001	CHERRY CREEK	Other	Xer	70.48	-2.63	1P, 2F

Site Code	SITENAME	RefStat	SiteClass	Tier 1 PCT_SAFN	Tier 2 LRBS_NOR	Assessment
CO022M	TEXAS CREEK	Strs	Mtn	51.52	-1.50	Fail
CO023M	SWIFT CREEK	Strs	Mtn	29.59	-1.08	1F, 2P
CO032M	CROOKED CREEK	Other	Mtn	10.10	-0.33	Pass
CO033		Other	Mtn	8.08	-0.30	Pass
CO034M	ALAMOSA RIVER	Other	Mtn	11.11	-0.64	Pass
CO035	ALAMOSA RIVER	Other	Mtn	16.67	-0.37	Pass
CO037M		Other	Xer	27.78	-0.31	Pass
CO038M	MIDDLE FORK NORTH CRESTONE	Ref	Mtn	4.08	-0.49	Pass
CO03RS		Ref	Mtn	13.19	-0.32	Pass
CO040M	JOHN'S CREEK	Other	Mtn	10.31	-0.85	Pass
CO04RS		Ref	Mtn	5.58	-0.24	Pass
CO04TS		Other	Mtn	0.85	0.01	Pass
CO054		Other	Mtn	43.01	-0.64	1F, 2P
CO056M		Strs	Mtn	16.37	-0.13	Pass
CO063M		Other	Mtn	7.07	0.09	Pass
CO066M	LOS PINOS	Ref	Mtn	21.21	0.42	1F, 2P
CO067M	Silver Creek	Ref	Mtn	3.03	-0.12	Pass
CO069M	CUNNINGHAM CREEK	Ref	Mtn	0.00	-0.78	Pass
CO06RS		Other	Mtn	3.37	0.45	Pass
CO06TS		Other	Mtn	9.89	0.03	Pass
CO070M	LIME CREEK	Ref	Mtn	1.01	0.28	Pass
CO072M	JUNCTION CREEK	Other	Mtn	5.05	-0.74	Pass
CO116M		Ref	Mtn	27.84	-1.12	Fail
CO117M		Ref	Mtn	17.17	-0.36	Pass
CO122M		Ref	Mtn	16.33	-0.41	Pass
CO127M		Other	Mtn	4.40	0.07	Pass
CO128M		Ref	Mtn	0.00	0.05	Pass
CO134M		Other	Mtn	1.01	-0.10	Pass
CO136M		Strs	Mtn	19.19	-0.29	Pass
CO162M		Ref	FtHill	23.47	-0.60	Pass
CO170M		Other	Mtn	7.07	-0.28	Pass
EPA01-0209	EAG1	Ref	FtHill	7.27	0.03	Pass
EPA01-0210	BLU4	Ref	FtHill	9.00	0.24	Pass
EPA01-0212	CMB1	Other	FtHill	60.95	-2.21	Fail
EPA01-0214	RUD1	Ref	FtHill	32.38	-1.92	1P, 2F
EPA01-0215	WLC1	Ref	Mtn	23.08	-0.76	1F, 2P
EPA01-0238	PURGATOIRE RIVER	Other	Xer	54.29	-1.91	Pass
EPA01-0239	SOUTH APACHE CREEK	Ref	Mtn	1.90	-0.83	Pass

Site Code	SITENAME	RefStat	SiteClass	Tier 1 PCT_SAFN	Tier 2 LRBS_NOR	Assessment
EPA01-0240	WAHATOYA CREEK	Ref	Mtn	2.61	-0.27	Pass
EPA01-0242	EL RITO AZUL	Ref	Mtn	0.00	-0.65	Pass
EPA01-0246	BIG BLUE CREEK	Ref	Mtn	13.33	-0.95	Pass
EPA01-0248	COTTONWOOD CREEK	Ref	Mtn	10.00	-0.84	Pass
EPA01-0302	S. FORK CAVE CREEK	Ref	FtHill	0.00	-0.68	Pass
EPA01-0305	FRYE CREEK	Ref	FtHill	1.90	-0.78	Pass
OWW04440-0045	SAN ANTONIO 2	Other	Mtn	19.05	-0.17	Pass
OWW04440-0077	CANONES CREEK	Other	Mtn	0.00	0.80	Pass
OWW04440-0205	Saladon Creek	Ref	Mtn	7.62	-0.12	Pass
OWW04440-0333	Rio Tusas	Other	FtHill	21.36	0.06	Pass
OWW04440-0429	Pecos River	Other	Xer	78.10	-3.39	Fail
OWW04440-0557	SAN ANTONIO (2)	Other	Mtn	34.26	-0.59	1F, 2P
OWW04440-0717	Rio Santa Barbara	Other	Mtn	0.95	0.43	Pass
OWW04440-0803	COMMISSION CREEK	toofar	Xer	82.86	-2.60	Fail
OWW04440-0845	Rio Nutrias	Other	FtHill	88.57	-2.23	Fail
OWW04440-1037	Negritos Creek	Ref	Mtn	25.71	-0.55	1F, 2P
OWW04440-1059	Canadian River	Strs	Xer	100.00	-1.23	1F, 2P
OWW04440-1069	JEMEZ CREEK	Other	FtHill	47.12	-1.80	Fail
OWW04440-1101	Wolf Creek	Other	Mtn	33.68	-0.82	1F, 2P
OWW04440-NM01	UTE Creek,NM	Ref	Xer	33.73	-1.71	Pass
OWW04440-NM03	Three Rivers	Ref	FtHill	18.10	-0.36	Pass
OWW04440-NM07	Turkey Creek	Ref	FtHill	33.65	-0.14	Pass
OWW04440-OK19	Cimarron River	Strs	Xer	100.00	-2.85	Fail
WAZP04-RBON1	BONITA CREEK	Ref	FtHill	24.55	-1.04	Pass
WAZP04-RLCR1	Little Colorado River	Ref	FtHill	1.74	0.08	Pass
WAZP04-RMIN1	MINERAL CREEK	Other	Mtn	14.78	-0.90	Pass

Site Code	SITENAME	RefStat	SiteClass	Tier 1 PCT_SAFN	Tier 2 LRBS_NOR	Assessment
WAZP99-0505		Strs	Mtn	3.81	-0.57	Pass
WAZP99-0512	Gila River	Strs	FtHill	81.90	-4.18	Fail
WAZP99-0537	Little Colorado River	Strs	Xer	100.00	-1.94	1F, 2P
WAZP99-0545	Black River	Ref	FtHill	19.05	-0.32	Pass
WAZP99-0569	KP CREEK	Ref	Mtn	4.76	-0.93	Pass
WAZP99-0599	Gila River	Strs	FtHill	60.00	-1.35	Fail
WAZP99-0605	Blue River	Strs	FtHill	18.10	-0.43	Pass
WAZP99-0615	CONKLIN CREEK	Other	Mtn	3.81	-0.36	Pass
WAZP99-0639	CAMPBELL BLUE CREEK	Other	FtHill	44.76	-2.04	Fail
WAZP99-0645	NUTRIOSO CREEK	Other	FtHill	86.67	-3.92	Fail
WAZP99-0648	Silver Creek	Other	FtHill	6.00	1.04	Pass
WAZP99-0653	NAZLINI CREEK	Other	Xer	33.33	-1.61	Pass
WAZP99-0669	TSAILE CREEK	Other	Mtn	35.24	-1.63	Fail
WAZP99-0681	Blue River	Ref	FtHill	22.86	-1.28	Pass
WAZP99-0687	CENTERFIRE CREEK	Ref	Mtn	22.86	-1.66	Fail
WAZP99-0701	BONITO CREEK	Other	FtHill	86.67	-2.43	Fail
WAZP99-0722	THOMPSON CREEK	Ref	Mtn	24.76	-1.70	Fail
WAZP99-0744	BUBBLING SPRING CANYON	Ref	Xer	99.05	-2.47	1F, 2P
WAZP99-0750	EAGLE CREEK	Ref	FtHill	23.81	-0.51	Pass
WAZP99-0768	BRIGHT ANGEL CREEK	Ref	Xer	23.81	-1.67	Pass
WAZP99-0783	LANPHIER CANYON	Ref	Mtn	12.38	-0.95	Pass
WAZP99-0828	NORTH FORK BLACK RIVER	Ref	Mtn	14.29	-0.38	Pass
WAZP99-0840	San Francisco River	Strs	Mtn	100.00	-2.99	Fail
WAZP99-0876	WHEATFIELDS CREEK	Strs	FtHill	91.43	-1.16	1F, 2P
WAZP99-0888	FISH CREEK	Ref	Mtn	3.81	-0.43	Pass
WAZP99-0906	Little Colorado River	Other	FtHill	58.10	-2.18	Fail
WCOP01-0734	SALT CREEK	Other	Xer	20.00	-1.08	Pass
WCOP01-0739	BIG DRY CREEK	toofar	Xer	52.38	-2.25	Pass
WCOP01-0752	WEST BIJOU CREEK	Other	Xer	100.00	-2.95	Fail
WCOP01-0765	WILD HORSE CREEK	Strs	Xer	87.37	-3.51	Fail
WCOP01-0777	CHACUACO CREEK	Ref	Xer	90.48	-2.35	1F, 2P
WCOP01-0809	WEST PLUM CREEK	Other	Xer	34.29	-1.22	Pass
WCOP01-0812	PURGATOIRE RIVER	Ref	Xer	36.89		1P
WCOP01-0817	MARKHAM ARROYO	Strs	Xer	93.14	-3.45	Fail
WCOP01-0819	TIMPAS CREEK	Other	Xer	45.71	-3.24	1P, 2F
WCOP01-0833	NORTH ST. CHARLES RIVER	Other	Mtn	40.91	-1.22	Fail

Site Code	SITENAME	RefStat	SiteClass	Tier 1 PCT_SAFN	Tier 2 LRBS_NOR	Assessment
WCOP01-0836	HORSE CREEK	Strs	Xer	85.44	-2.25	1F, 2P
WCOP03-R001	SOUTH RUSH CREEK	toofar	Xer	99.04	-2.87	Fail
WCOP03-R005	AGATE CREEK	Other	Mtn	29.52	-1.12	Fail
WCOP03-R007	EAST FORK HERMOSA CREEK	Other	Mtn	38.10	-1.04	1F, 2P
WCOP03-R008	Bear Creek	Ref	Mtn	5.71	-0.57	Pass
WCOP03-R009	EAST FORK PIEDRA RIVER	Ref	Mtn	7.62	-0.60	Pass
WCOP04-R003	TWO BUTTE CREEK	Ref	Xer	47.62	-1.52	Pass
WCOP04-R007	YELLOW JACKET CREEK	Ref	Xer	43.81	-1.88	Pass
WCOP04-R009	TIMPAS CREEK	Ref	Xer	67.62	-2.92	1P, 2F
WCOP99-0502	ADAMS FORK CONEJOS RIVER	Ref	Mtn	4.55	-0.91	Pass
WCOP99-0507	GROUNDHOG CREEK	Other	Mtn	0.95	-0.49	Pass
WCOP99-0508	RED MOUNTAIN CREEK	Ref	Mtn	6.67	-0.83	Pass
WCOP99-0509		Other	Xer	60.00		1P
WCOP99-0510	Wolf Creek	Strs	Xer	63.81	-2.23	Pass
WCOP99-0513	WHITEHOUSE CREEK	Ref	Mtn	13.33	-1.22	1P, 2F
WCOP99-0563	SOUTH RUSH CREEK	Other	Xer	94.29	-2.34	1F, 2P
WCOP99-0568	LA PLATA RIVER	Other	Xer	30.48	-1.21	Pass
WCOP99-0569	EAST PLUM CREEK	toofar	Xer	41.76	-1.47	Pass
WCOP99-0574	HENSON CREEK	Other	Mtn	23.08	-1.07	1F, 2P
WCOP99-0591	FALL CREEK	Other	Mtn	11.43	-0.47	Pass
WCOP99-0622	HARTMAN DRAW	Strs	Xer	61.17	-1.54	Pass
WCOP99-0627	HOUSELOG CREEK	Other	Mtn	64.76	-1.93	Fail
WCOP99-0629	WEST PLUM CREEK	Other	Xer	63.81	-1.52	Pass
WCOP99-0634	Ute Creek	Ref	Mtn	14.29	-1.66	1P, 2F
WCOP99-0646	MUD CREEK	Strs	Xer	48.57	-2.12	Pass
WCOP99-0670	LOST CANYON CREEK	Other	Mtn	12.38	-0.48	Pass
WCOP99-0672	PURGATOIRE RIVER	Other	Xer	41.62	-1.86	Pass