A Statewide Assessment of BLM-Managed Streams and Rivers in Idaho

Scott W. Miller[1], Nicole Cappuccio[1], Jennifer Courtwright[1], Lucy Littlejohn[2], Bryce Bohn[3]

[1]BLM National Operations Center, Branch of Assessment and Monitoring, Lakewood, Colorado
[2]BLM Salmon Field Office, Fisheries and Aquatic Habitat Program, Salmon, Idaho (Retired)
[3]BLM State Office, Fisheries and Aquatic Habitat Program, Boise, Idaho (Retired)

Table of Contents

Summary of Findings

The Bureau of Land Management's (BLM's) lotic assessment, inventory, and monitoring (AIM) group, in collaboration with its Aquatic Habitat Management staff in Idaho, compiled multiple AIM National Aquatic Monitoring Framework surveys to report on the condition of BLM-managed wadeable streams and rivers in Idaho. We used chemical, physical, and biological data from 65 randomly located BLM reaches sampled between 2013 and 2016. Overall, we found BLM-managed streams and rivers in Idaho to be in moderate to good condition, with the greatest concerns related to nutrient loading, streambank conditions, and biological condition as indicated by macroinvertebrates. For example, nearly one-third of perennial stream kilometers have degraded biological condition, and excessive nutrient loading is a widespread contributor to observed biological conditions. In contrast, more than one-third of BLM-managed streams had less bank stability and cover and nearly half had less canopy cover than their desired condition, but neither condition significantly impacted biological condition. Lastly, pH exceedances, while localized in nature, had the most significant impacts to biological condition when present. As a follow up to this assessment, BLM's Aquatic Habitat Management staff in Idaho must track changes in chemical, physical, and biological conditions by repeating this survey to identify trends and determine management effectiveness over time. The Idaho Falls, Twin Falls, and Coeur d'Alene Districts are working to do this by assessing lotic habitat conditions at the field office and project scales. Idaho BLM staff are also working with the AIM group to identify land uses associated with priority stressors and degraded biological condition to determine whether BLM-permitted activities are causal factors in the observed departures from land health standards. In addition to assessing lotic habitat conditions at the statewide scale, the

Background

The BLM oversees approximately 47,000 square kilometers (km^2) of land containing over 6,085 kilometers (km) of perennial streams and rivers throughout the state of Idaho. BLM stream and riparian systems are among the most important, productive, diverse, and sensitive ecosystems in the state. They support a wide range of aquatic and terrestrial species and provide water for humans, wildlife, and livestock as well as irrigated agriculture and other ecosystem services.

Under the Federal Land Policy and Management Act (FLPMA), the BLM manages the public lands, now known as the National System of Public Lands, for multiple uses and for sustained yield. Thus, watersheds are managed both for conservation of natural resources and for activities that use and impact riparian and aquatic resources, such as livestock grazing, timber harvest, mining, energy development, transportation networks, and dispersed and developed recreation activities. Consequently, knowing the condition and trend of riparian and aquatic ecosystems is critical to achieving the BLM's mission of "sustaining the health, diversity, and productivity of public lands for the use and enjoyment of present and future generations."

Section 201(a) of FLPMA (43 U.S.C. 1711(a)) requires the periodic and systematic inventory of renewable resource condition and trend. As part of its assessment, inventory, and monitoring (AIM) strategy, the BLM is implementing the National Aquatic Monitoring Framework (NAMF) to provide baseline condition and monitoring data. These data inform decisions at multiple scales, such as individual restoration and reclamation projects and permitted uses at the local scale, resource management plan effectiveness at the regional scale, and aquatic resource condition reporting at the national scale (BLM 2015). For this effort, we compiled multiple past and ongoing AIM NAMF surveys to report on the condition of wadeable streams and rivers (hereafter referred to as streams) managed by the BLM in Idaho. Specifically, we used information from the BLM's Western Rivers and Streams Assessment (2013–2015) and various intensification efforts that occurred in 2016 throughout Idaho.

We designed this statewide survey of chemical, physical, and biological conditions on BLM-managed wadeable streams in Idaho to achieve three main objectives:

1. Determine baseline chemical, physical, and biological conditions for Idaho-managed stream systems from which trends can be assessed.
2. Assess attainment of Idaho standards for rangeland health for stream systems.
3. Identify and rank stressors contributing to degraded stream conditions.

Methods

In total, we used 65 stream reaches from the target population of wadeable streams occurring on BLM-managed lands in Idaho (figure 1). All points were from probability-based survey designs where points were originally allocated in proportion to the linear extent of streams by Strahler stream order categories (small streams: first and second order; large streams: third and fourth order; and rivers: fifth+ order) (Strahler 1952). The random nature of the points and use of identical field methods among studies allowed us to combine points to report on the condition of BLM-managed streams statewide with minimal bias and known levels of confidence.

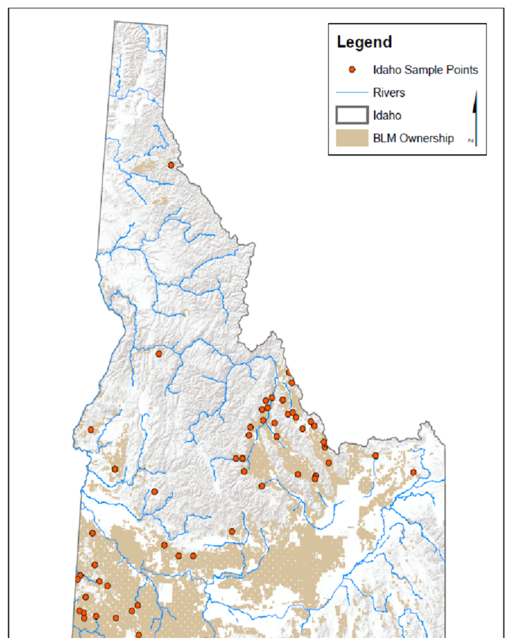


Figure 1. Distribution of the 65 wadeable stream and river reaches sampled between 2013 and 2016 throughout BLM-managed lands in Idaho.

The 65 reaches were sampled during the summers of 2013–2016 using AIM methods (BLM 2017), which collectively address four of the Idaho BLM land health standards applicable to stream systems: *Standard 2*: Riparian Areas and Wetlands; *Standard 3*: Stream Channel/Floodplain; *Standard 7*: Water Quality; and *Standard 8*: Threatened and Endangered Plants and Animals (U.S. DOI BLM 1997). From the field data, we report on four indicators of stream channel and riparian function (Idaho standards 2, 3, and 8) and six indicators of water quality (Idaho standard 7).

To assess stream conditions objectively, we established benchmarks for each of the 10 indicators (table 1). Benchmarks are indicator values used to assess whether a reach is meeting or not meeting standards based on how observed conditions compare to reach potential in the absence of anthropogenic constraints. Where policy did not specify benchmarks, we used existing monitoring networks that have quantified least disturbed "reference" reaches to establish the Idaho benchmarks (i.e., desired conditions based on reach potential). The reference network used depended on the indicator, as no single network encompassed all indicators. Our analysis used reference networks compiled by Olson and Hawkins (2012, 2013) for total nitrogen, total phosphorus, and specific conductance; Herlihy et al. (2008), using Environmental Protection Agency (EPA) stream assessment data, for canopy cover, floodplain connectivity, and percent fine sediment; and U.S. Forest Service/BLM PACFISH/INFISH Biological Opinion (PIBO) (Irvine et al. 2014) for macroinvertebrates. See appendix A for more detailed descriptions of the reference networks and methods for determining benchmarks. All approaches for quantifying reach potential sought to minimize the presence of human impacts, as indicated by land uses and surface disturbances, such as percent of agriculture or urban land use, road density, timber harvest, and grazing. Overall, the PIBO reference network had a more stringent set of criteria than both Olson and Hawkins and the EPA. Regardless of the reference network used, we used our best professional judgement to ensure the benchmarks made ecologic, hydrologic, or geomorphic sense.

The analysis established benchmarks in one of two ways using reference networks. The first method used empirical models to make reach-specific predictions of the conditions expected to occur in the absence of anthropogenic impairment, referred to as predicted natural conditions (PNC). The alternative method quantified the range of variability among reference reaches by ecoregion and stream size. For either approach, the analysis asked whether observed conditions were within reach-specific predictions or reference reach variability (meeting the benchmark) or at the extremes (not meeting the benchmark) to make condition determinations.

We report results in terms of the extent of BLM-managed wadeable streams in Idaho meeting or not meeting specified benchmarks with a ±90% confidence interval (i.e., relative extent). Such results allow managers to determine how pervasive a stressor or problem is throughout the state (i.e., relative spatial extent of one indicator versus another). Because we randomly selected sample reaches, we can make inferences to all BLM-managed wadeable streams in Idaho with known levels of precision and accuracy.

We also reported "relative risk" as a measure of potential impacts of each indicator to macroinvertebrate biological condition to complement indicator relative extent estimates. Relative risk values provide insight into what the presence of a stressor means to one of the beneficial uses identified in the State of Idaho's water quality standards, which for many BLM-managed streams includes aquatic life support under the Clean Water Act. Relative risk compares the likelihood of not meeting the macroinvertebrate biological condition benchmark when a given stressor is also not meeting a benchmark to the likelihood of not meeting the macroinvertebrate biological condition benchmark when a stressor is meeting a benchmark (Van Sickle et al. 2006). The ratio of these two likelihoods is the relative risk.

For example, a relative risk value of 1.5 for total phosphorus means that a stream reach is 1.5 times more likely to not meet the macroinvertebrate biological condition benchmark when a reach also does not meet the total phosphorus benchmark. In contrast, a relative risk of 1.0 means that the reach is equally likely to not meet the macroinvertebrate biological condition benchmark when stressor levels are high as when they are low. From a stressor standpoint, relative risk values greater than 1.0 are of greatest concern, especially where the lower extent of the 90% confidence interval is also greater than 1.0.

Table 1. Field methods and condition benchmarks for the subset of 10 AIM indicators used to report on the condition of BLM-managed wadeable streams in Idaho. See appendix A for more detailed descriptions of the methods used to determine benchmarks.

Indicator	Summary field method	Condition benchmark
Total nitrogen	Single grab sample for lab analysis	Observed value ≤ PNC plus 95th percentile of model error (114.7 µg/L) [1]
Total phosphorus	Single grab sample for lab analysis	Observed value ≤ PNC plus 95th percentile of model error (21.3 µg/L) [1]
Specific conductance	In situ: multiparameter sonde	Observed value ≤ PNC plus 95th percentile of model error (74.5 µS/cm) [1]
pH	In situ: multiparameter sonde	Observed value ≥6.5 standard units (SU) and ≤9 SU [2]
Water temperature	NA – mean August stream temperature derived from NorWesST (e.g., Isaak et al. 2017)	Predicted value ≤ Idaho Department of Environmental Quality (DEQ) standard of 19 °C for cold-water systems [3]
Macroinvertebrates	8 comporeach Surber samples from riffle habitats or multihabitat sampling consisting of 11 comporeach Surbers	Observed value ≥0.63 based on PIBO observed/expected (O/E) macroinvertebrate model
Percent fine sediment (< 2 mm)	10 particles from each of 21 transects from active channel	N. Rockies: Observed value ≤29% for small streams and 15% for large streams; [4] N. Xeric Basins: Observed value ≤45% for small streams and 44% for large streams [4]
Bank cover and stability	Left and right bank at 21 transects	Observed value ≥80% for both bank stability and cover based on Idaho BLM policy
Percent bank canopy cover	Left and right bank at 11 transects with densiometer	N. Rockies: Observed value ≥76.5% for small streams and 61% for large streams; [4] N. Xeric Basins: Observed value ≥69% for small streams and 55% for large streams [4]
Channel incision (only on streams ≤4% slope)	Bankfull and floodplain heights at 11 transects	N. Rockies: Observed value ≤0 for small streams and 0.01 for large streams; N. Xeric Basins: Observed value ≤-0.09 for small streams and 0.11 for large streams [4]

[1]PNC derived from models of Olson and Hawkins (2012, 2013)

[2]pH benchmarks based on Idaho DEQ and EPA guidance

[3]Temperature benchmark based on Idaho DEQ guidance for the protection and maintenance of a viable aquatic life community for cold-water species

[4]Benchmarks for fine sediment, canopy cover, and channel incision are based on the percentiles of regional reference conditions following Kaufmann et al. (1999) and Stoddard et al. (2005). Percentiles computed for aggregate level III ecoregions including the Northern Rockies and Northern Xeric Basins for streams less than 10 m (small) and greater than 10 m (large) in bankfull width.

Results and Discussion

Inventory of perennial streams and rivers

Our original estimate of the extent of BLM-managed wadeable, perennial streams in Idaho was 6,085 km, based on the U.S. Geological Survey's (USGS's) National Hydrography Dataset (NHD) Plus v2. However, after reach scouting and sampling, we found that 38% of those streams (2,300 km) were nonperennial, and an additional 13% (773 km) were not within our target group for other reasons (e.g., not on BLM-managed land, irrigation canal, wetland) (figure 2). Ultimately, during the period of the study (2013–2016) we estimated the extent of BLM-managed wadeable, perennial streams in Idaho to be 3,012 km. Within those stream kilometers, we used 65 randomly placed sample reaches to make inference to 2,395 stream kilometers, attributing the difference (3,012 km – 2,395 km) to access issues.

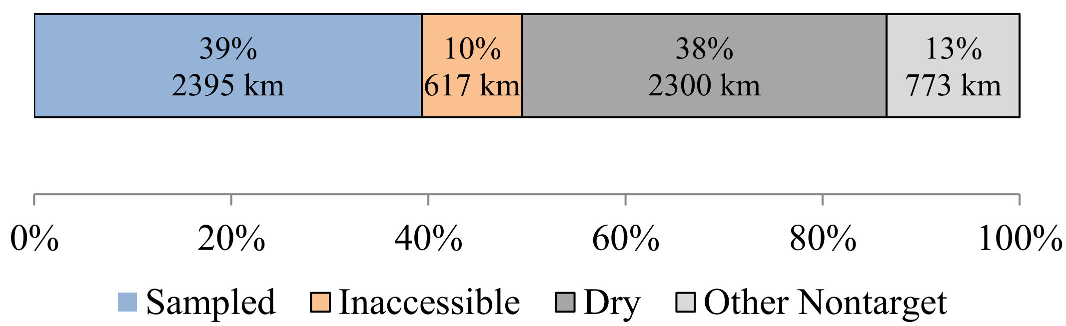


Figure 2. Percent of perennial stream kilometers from the National Hydrography Dataset that were sampled, inaccessible, and determined not to be within our target group (e.g., dry, not on BLM-managed land, and other reasons).

We attempted to determine why such a large extent (38%) of streams classified as perennial in the National Hydrography Dataset (NHD) were dry. Specifically, we tried to determine whether relationships existed between observed dry reaches and anthropogenic activities such as dams, irrigation diversions, and agricultural land uses. However, we found no significant relationships (results not shown). Rather, we found that dry reaches were more likely to occur in response to natural physiographic conditions such as watersheds with small areas and low precipitation (results not shown). New and ongoing AIM projects at the district and field office scales will need to verify where the perennial streams are located within the study area and identify whether the impacts of stream dewatering are a locally relevant management concern.

Water quality indicators

Water quality conditions were generally good across the state, apart from potential nutrient exceedances and degraded biological assemblages. For example, few streams had pH, specific conductance, or temperature exceedances, with over 89% of stream kilometers characterized as meeting management benchmarks for each of these indicators (figure 3, Appendix B). In contrast, excessive nutrients and degraded macroinvertebrate biological condition were far more widespread. For example, only about one-half (52%) of stream kilometers met the benchmark for total nitrogen and almost two-thirds (64%) of the stream kilometers met the benchmark for total phosphorus. Macroinvertebrate biological condition showed the most departure from reference, with just over one-third (36%) of stream kilometers meeting the management benchmark (figure 4, Appendix B).

Although Idaho state water quality standards do not exist for specific conductance, total nitrogen, and total phosphorus, we used models to assess the extent to which observed concentrations exceeded PNC (i.e., reach potential). For nutrients, a significant proportion (36%–48%) of BLM-managed stream kilometers in Idaho exceeded PNC. Given the lack of state numeric criteria for nutrients, such exceedances are not in violation of water quality standards. However, the relative risk results show that degraded macroinvertebrate biological condition was 2.0 times more likely to occur when the total nitrogen benchmark was exceeded and 1.5 times more likely to occur during phosphorus exceedances (figure 6). Consequently, the observed nutrient exceedances were both pervasive and associated with adverse impacts to stream biota. We based our analyses on one-time grab samples collected during summer base-flow conditions and, therefore, these observations require validation through more temporally intensive sampling. However, the large number of reaches with elevated nutrient levels suggests that our results are not likely unusual.

Management activities that accelerate erosion rates, change biogeochemical processes, or directly add nutrients to streams, such as logging, cattle grazing, and agriculture, can significantly increase nutrient loading to streams (reviewed in Allan and Castillo 1995). Eutrophication can have adverse impacts to water quality conditions, such as lowered dissolved oxygen levels resulting from increased rates of primary production and subsequent organic matter decomposition. Eutrophication can also indirectly impact biological assemblages, such as macroinvertebrates and fishes, through changes to the trophic basis for secondary production and subsequent food web structure as well as through alterations to physical habitat, such as water clarity and benthic substrates (Dodds 2006; Miller and Crowl 2006; Dunck et al. 2015).

The study found that excessively high or low pH values were only a problem in 5.7% of BLM-managed wadeable streams in Idaho (figure 3, Appendix B). However, when we observed pH levels that were out of balance, they significantly impacted macroinvertebrate biological conditions. For example, failure to meet the macroinvertebrate biological condition benchmark was 2.7 times more likely in streams with pH irregularities (figure 6). Such conditions were rare and most likely related to isolated mining activities (low pH) and high rates of photosynthesis (high pH) or respiration (low pH) in the presence of aquatic macrophytes or high algal densities (Tank et al. 2009; Hogsden and Harding 2012). We observed similar patterns for specific conductance, with only 10.7% of streams impacted, but when specific conductance exceeded the benchmark, degraded macroinvertebrate biological condition was 1.8 times more likely.

Finally, due to the large geographic scope of this report, we based assessments of stream temperature on how predicted values across all BLM-managed lands in Idaho compared to the Idaho cold-water beneficial use benchmark of 19 ˚C (Table 1). The relatively high temperature benchmark relative to the thermal requirements of bull trout (13 ˚C) or other spawning salmonids (9–13 ˚C) likely explains why we observed only a small proportion of streams exceeding the benchmark. The 8% of stream kilometers that did exceed the benchmark were largely limited to larger order systems throughout the Snake River Plain (figure 7). If we applied lower water temperature benchmarks, we would likely see a much larger proportion of systems not meeting the standard, including smaller order systems critical to fish spawning and juvenile rearing.

WATER QUALITY INDICATORS

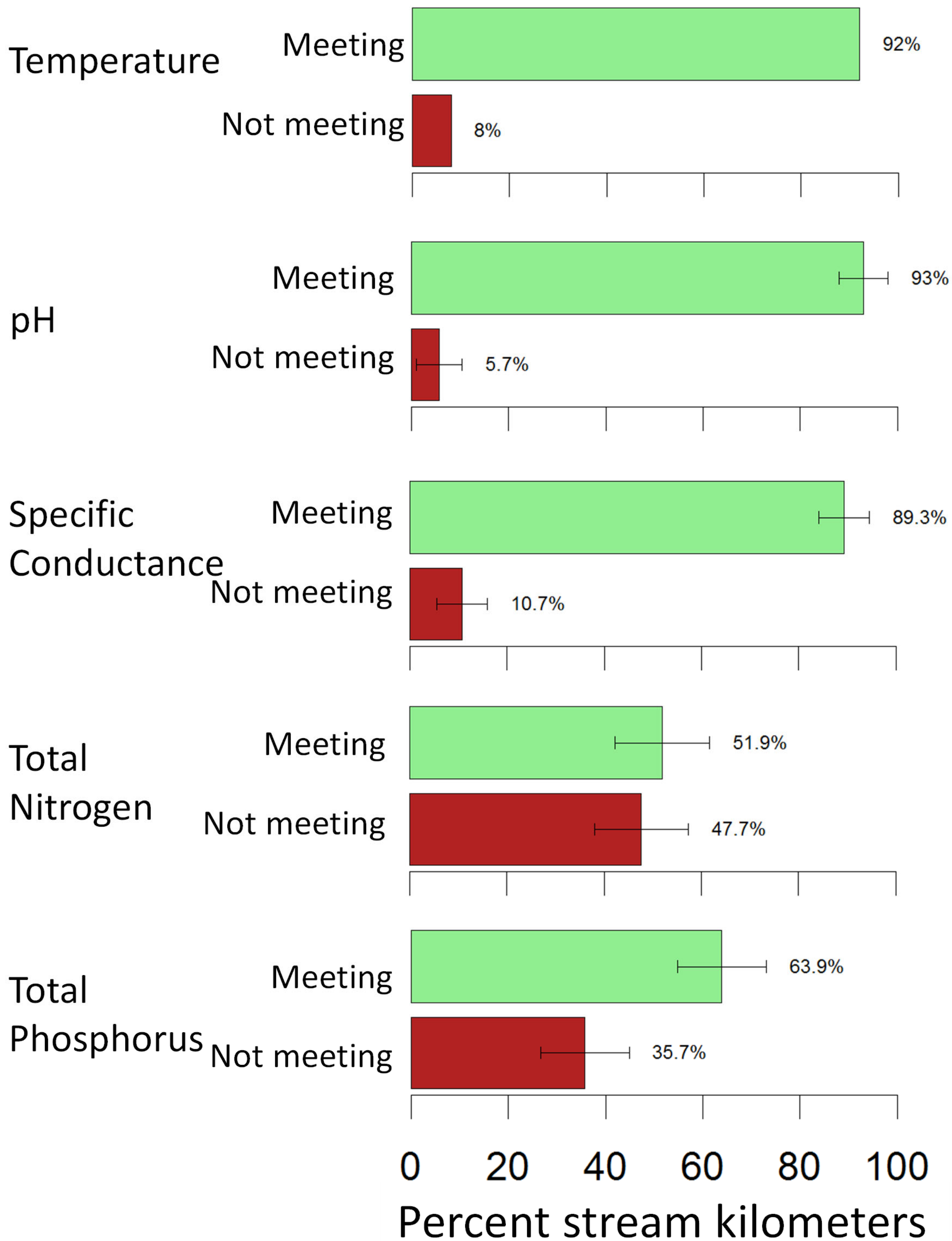


Figure 3. Relative extent of stream kilometers meeting versus not meeting benchmarks for the assessed water quality indicators (±90% confidence intervals). Of the assessed indicators, we based only pH and temperature conditions on state standards (see table 1 and appendix A for

details). For some indicators, the two condition categories do not sum to 100% because of missing data.

BIOLOGICAL CONDITION INDICATOR

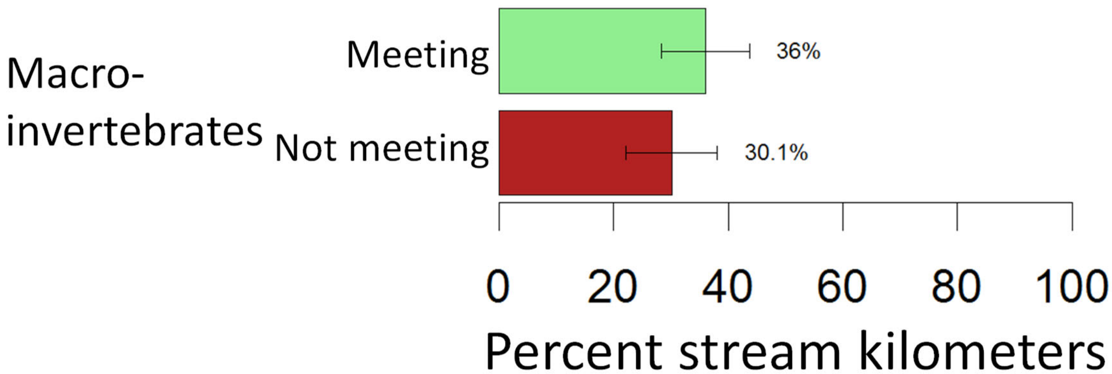


Figure 4. Relative extent of stream kilometers meeting versus not meeting the benchmark for macroinvertebrate biological condition (±90% confidence intervals). The two condition categories do not sum to 100% because of low macroinvertebrate counts observed at some reaches or reaches falling outside of the environmental conditions of reference reaches used to develop the O/E model. Low counts (e.g., <200 individuals) can influence macroinvertebrate O/E scores and can result from degraded biological condition, sampling error, or a combination of the two. Determining whether these conditions are real or are a sampling artifact will require resampling of the reaches with low counts.

Riparian and instream function indicators

Of the four indicators used to assess the physical condition of riparian and stream conditions, percent fine sediment was in the best condition, with nearly 80% of assessed BLM-managed stream kilometers in Idaho meeting the benchmark (figure 5, Appendix B). Bank stability, cover, and percent canopy cover along streambanks varied more. Banks that were both stable and covered had 63% of stream kilometers meeting the cover or stability benchmark. Canopy cover had 55% of stream kilometers meeting the management benchmarks across ecoregions and stream sizes (figures 5, B4, B6). Lastly, 40% of assessed stream kilometers met the channel incision benchmark, while we considered 9% incised and disconnected from the floodplain.

In contrast to some water quality exceedances, the presence of instream habitat stressors such as fine sediment only moderately increased the relative risk of not meeting the macroinvertebrate biological condition benchmark (figure 6). Similarly, riparian conditions such as channel incision were not associated with the likelihood of observing major departure in biological condition.

There are two important caveats for the riparian and instream function indicators. First, we only assessed channel incision for streams with the capacity to support a floodplain, defined as reaches with a slope less than 4% and this equated to 58.3% of BLM-managed stream kilometers. Also, the channel incision assessments only evaluated geomorphic and not hydrologic conditions. Secondly, we included all data for fine sediment regardless of slope, which meant the inclusion of both transport and depositional reaches. The inclusion of all data did not appear to overly skew the fine sediment results, however, as we only observed a weak relationship between slope and fine sediment (R^2 <10%) and only 19% of the low-gradient reaches (slope <4%) exceeded the benchmark compared to 20.5% of reaches overall.

The assessment of bank cover and stability evaluates the susceptibility of streambanks to both natural and accelerated erosion rates associated with land management activities. Anthropogenic activities that increase stream power (e.g., flow alteration, improperly sized culverts) or alter the composition and cover of stabilizing vegetation (e.g., grazing, recreation, invasive species) can increase bank erosion rates (Booth 1990; Knapp and Matthews 1996; Zaimes et al. 2004; Herbst et al. 2012). Streambank erosion can increase fine sediment loading to streams and reduce the viability of the stream bottom environment for aquatic organisms such as amphibians, macroinvertebrates, and fishes (Cunjak and Power 1986; Bjornn and Reiser 1991; Wood and Armitage 1997; Henley et al. 2000). For example, we found that failure to meet the macroinvertebrate biological condition benchmark was 1.8 times more likely in streams not meeting the fine sediment benchmark (figure 6). Excessive bank erosion can also alter channel morphology and subsequent habitat quality (e.g., width:depth ratios) through altering the balance between the sediment and water supply (Lyons et al. 2000; Ripple and Beschta 2006).

The assessment of percent canopy cover evaluates the capacity of riparian vegetation to mitigate thermal loading (i.e., provide shade) and thus moderate stream temperatures (Beschta 1997; Johnson and Jones 2000). The extent of the riparian canopy also provides information on the amount of potential leaf litter to subsidize aquatic food webs (Cummins 1974). Alterations to canopy cover were fairly widespread (45% of Idaho BLM streams have canopy cover less than potential) and are potentially impacting biological condition; reaches were 1.6 times more likely to not meet the macroinvertebrate biological condition benchmark when the reach is also not meeting the canopy cover benchmark (figure 6). In contrast, stream temperature exceedances were not an issue based on the benchmark of 19 °C; an assessment at the field office scale should use more stringent temperature benchmarks based on resident fish species.

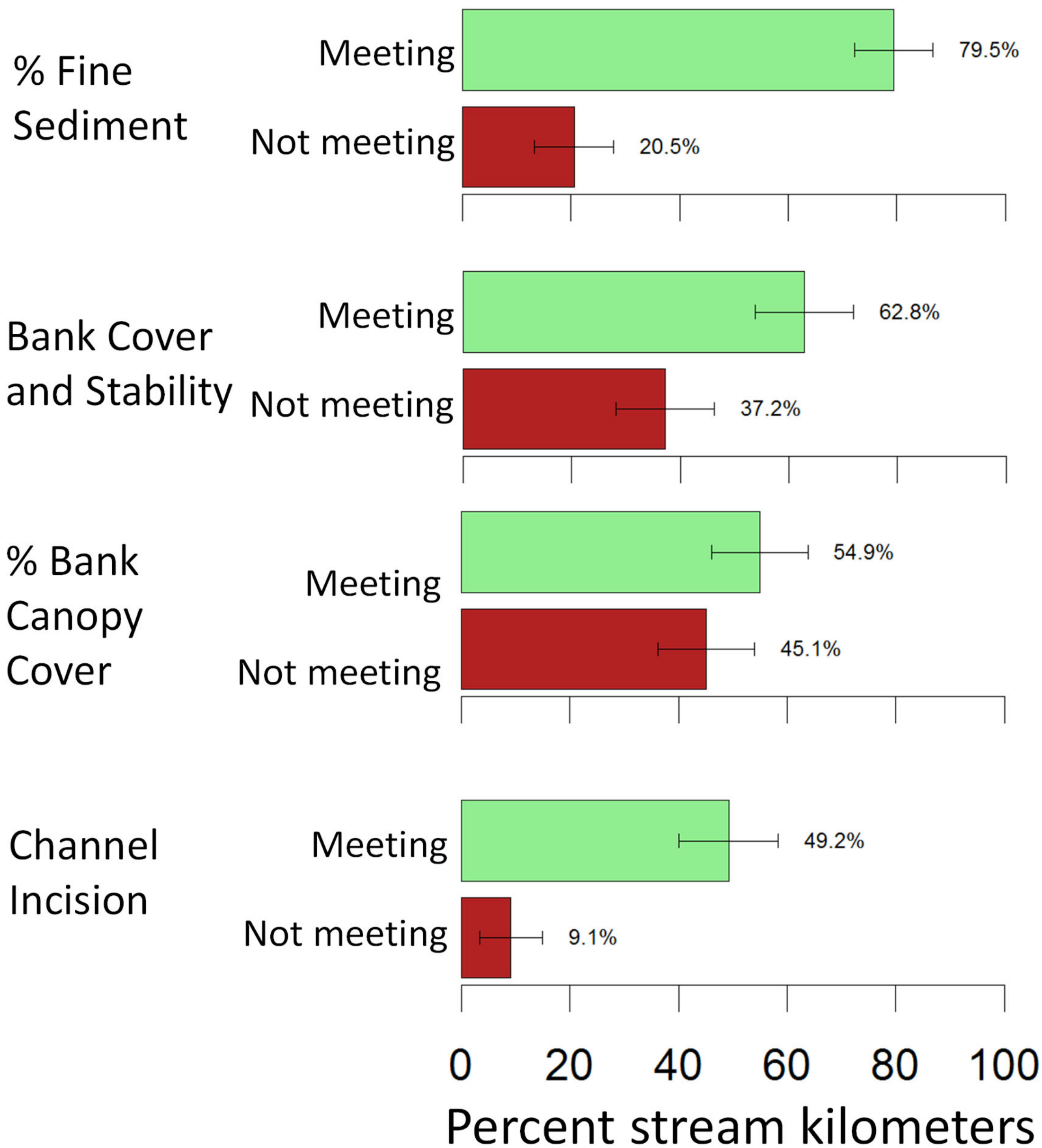


Figure 5. Relative extent of stream kilometers meeting versus not meeting benchmarks for the assessed riparian and instream habitat indicators (± 90% confidence intervals). Of the assessed indicators, we based only bank cover and stability conditions on state standards (see table 1 and appendix A for details). For some indicators, the two condition categories do not sum to 100% because of missing data or the exclusion of high-gradient reaches in the case of channel incision.

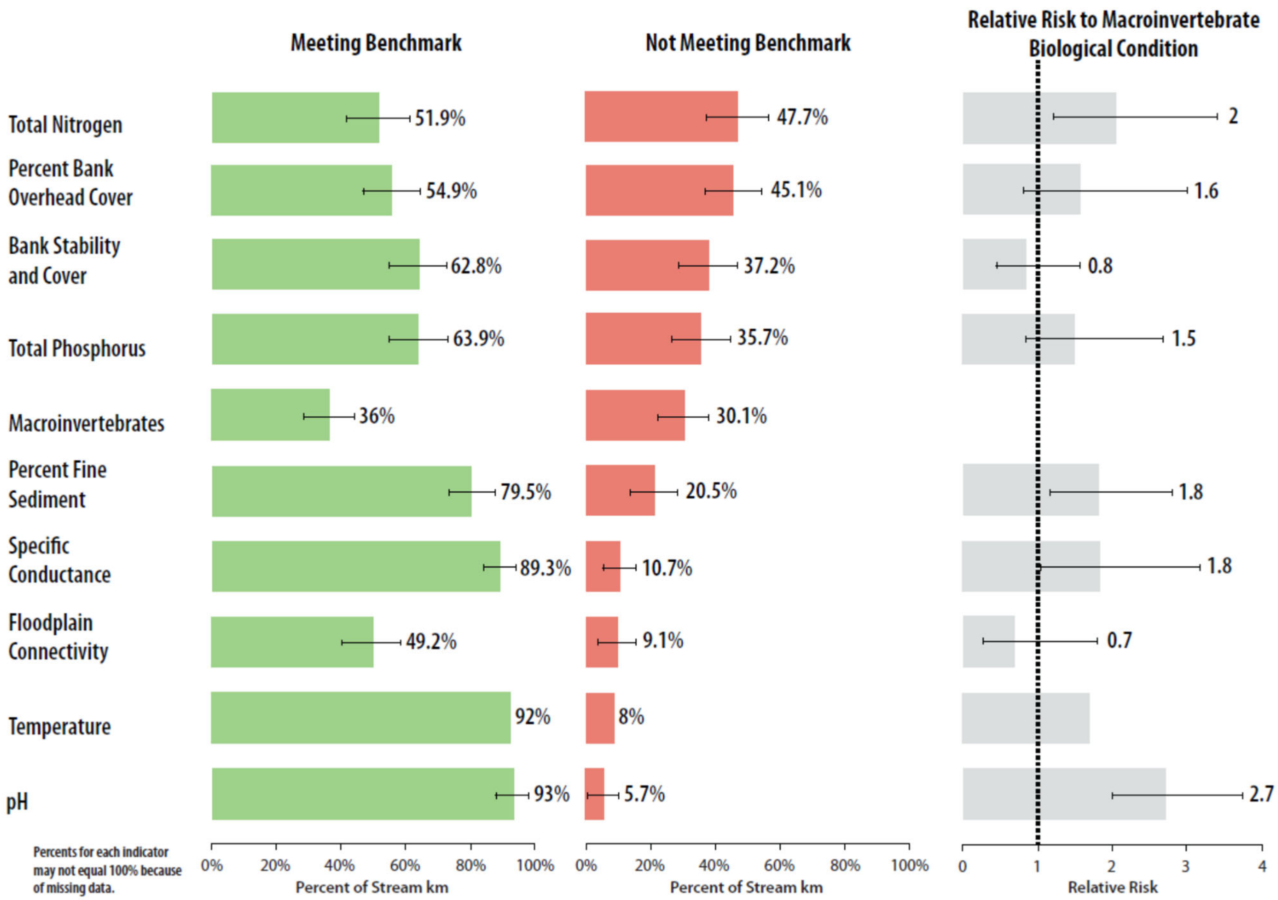


Figure 6. Relative extent of water quality, riparian, and instream indicators (i.e., stressors) meeting and not meeting the management benchmarks specified in table 1 and the relative risk of those stressors to the biological condition of macroinvertebrate assemblages (± 90%). For example, a reach is two times more likely to not meet the macroinvertebrate biological condition benchmark when total nitrogen also does not meet the specified benchmark.

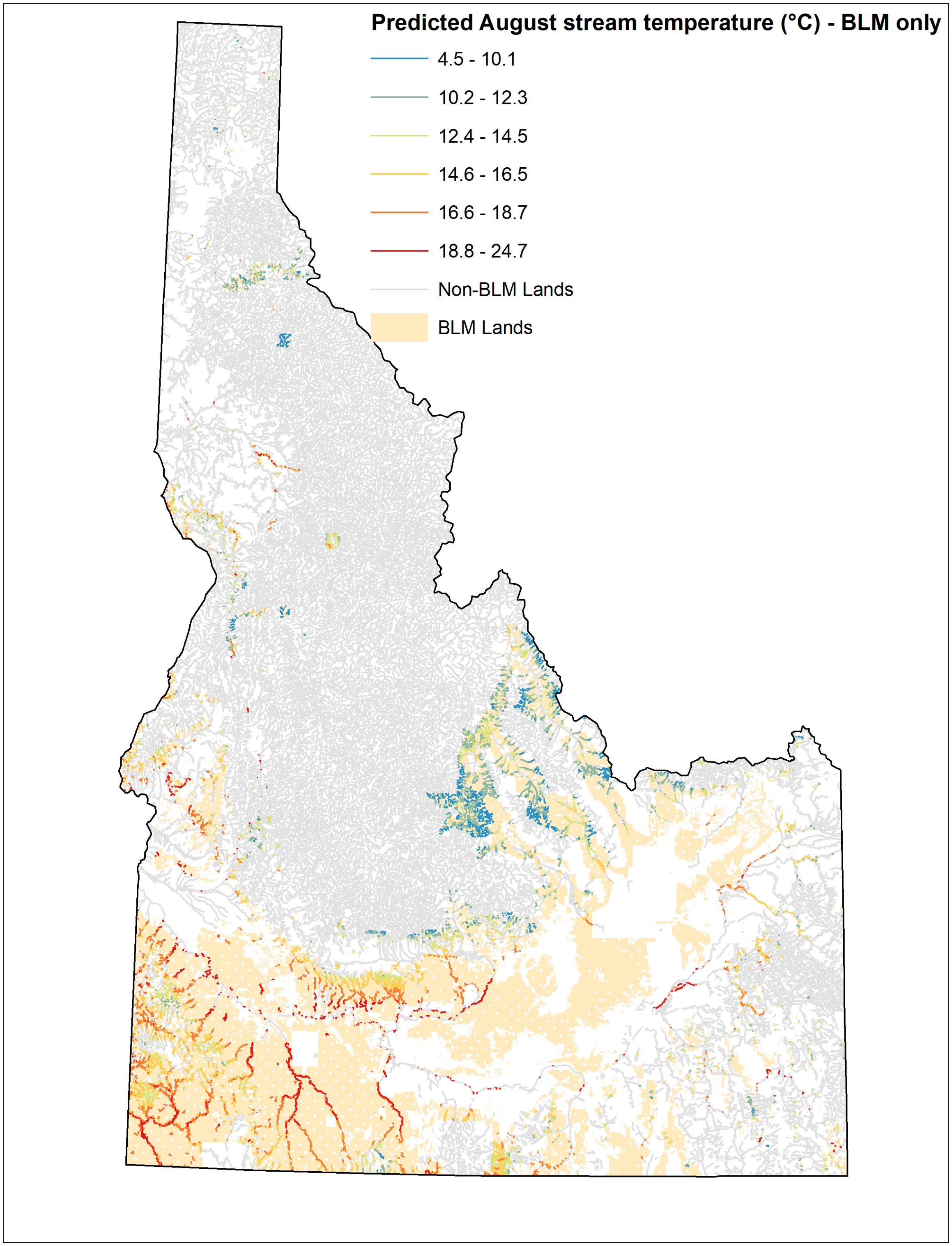


Figure 7. Predicted stream temperature for BLM-managed streams in Idaho. While stricter criteria exist for fishes such as bull trout, we used a benchmark of 19 °C following Idaho DEQ guidance for the protection and maintenance of a viable aquatic life community for cold-water species.

Management Priorities and Next Steps

Now that the statewide lotic AIM assessment is complete, the next step is to identify whether land uses are associated with priority stressors and degraded biological condition, and to determine whether and to what extent BLM permitted activities may be causal factors. Examples include determining whether BLM-managed streams are inheriting excessive nutrient loading from upstream landowners or whether activities such as livestock grazing or excessive upland erosion on BLM-managed land are contributing to observed conditions. The BLM will use such information to identify best management practices, strengthen collaborations with state and federal partners to improve watershed health, and ensure the productivity and sustainability of BLM-managed rangelands and permitted activities.

To that end, the BLM is currently implementing lotic AIM monitoring projects at the field office scale throughout the Idaho Falls, Twin Falls, and Coeur d'Alene Districts. The objectives of these efforts are multifaceted but include assessing the effectiveness of resource management plans, providing data to inform land health determinations, and assessing restoration efficacy. Data resulting from these projects are compatible with the statewide assessment and the BLM will use these data to better understand causes for observed conditions.

An unanticipated outcome of this assessment was the discovery of the misclassification of nearly 40% of BLM-managed streams in the NHD, which suggests a need for the BLM to work with the USGS and others to refine the NHD stream layer. Accurate inventories of these types of lotic systems (i.e., intermittent or perennial) are essential to all BLM planning and inventory, assessment, and monitoring activities. This issue is of particular concern given climate change, recent droughts, and management actions, which could change the flow periodicity of BLM-managed streams. In addition to improving the quality of the NHD, the BLM should expend effort to monitor changes in flow periodicity among systems.

Lastly, an important element of AIM and the BLM's assessment efforts is the ability to track change in the chemical, physical, and biological condition of streams through time. This report provides a quantitative, unbiased baseline from which the BLM can track the cumulative effectiveness of management actions over time. The BLM will repeat starting in 2021.

Appendix A: Description of Reference Reach Networks and Methods Used for Benchmark Development

Overview of PACFISH/INFISH Biological Opinion macroinvertebrate observed/expected model and reference criteria

We used the U.S. Forest Service/BLM PACFISH/INFISH Biological Opinion (PIBO) observed/expected (O/E) index to assess biological condition of sampled reaches based on the idea of predicted natural conditions (PNC). O/E models compare the macroinvertebrate taxa observed (O) at reaches of unknown biological condition (i.e., "sampled reaches" or "test reaches") to the assemblages expected (E) to be found in the absence of anthropogenic stressors (see Hawkins et al. 2000 for details). The PIBO O/E model is based on 201 reference reaches (174 calibration and 27 randomly selected validation reaches) grouped into 10 distinct classes based on the similarity of macroinvertebrate assemblage composition among reaches following the standard methods of Hawkins et al. (2000) and described in detail by Stoddard, Peck, Olsen, et al. (2005). The expected class membership and subsequent reference macroinvertebrate assemblage (E) for comparison to test reaches is predicted by linear discriminant function models using the following predictors: watershed area, 30-year average monthly maximum air temperature, and 30-year average precipitation for the 12 months prior to a standardized sample collection date of July, both derived from PRISM estimates. Prior to computing O/E scores, we standardized data for all test reaches to the operational taxonomic units and resampled to a fixed count of 300. Following the recommendations of Hawkins et al. 2000, we calculated O/E scores for taxa having a probability of capture ≥0.5 to increase the precision of O/E estimates and subsequent model sensitivity to stressors.

We subsequently assessed biological condition based on the precision of the 27 validation reaches (mean = 0.95, standard deviation [SD] = 0.16), with test reaches scoring less than 2 SDs below the mean of reference reaches "meeting" the benchmark (i.e., comparable to reference conditions) and reaches scoring more than 2 SDs below the mean of reference reaches "not meeting" the benchmark. For the PIBO O/E model, the minimum count required for computing an O/E score is 200 individuals. Samples with less than 200 individuals do not get a condition rating. Lastly, prior to applying the PIBO model, we used Mahalanobis distance following Hawkins et al. (2000) to assess whether the environmental conditions for each of the 65 sampled reaches were typical of the 174 reference reaches used to develop the PIBO O/E model. In the end, we only used 43 of the 65 collected samples to assess biological condition.

The 201 PIBO reference reaches are designated by PIBO based on management history. Specifically, reference reaches are primarily located in wilderness areas or in subwatersheds with no obvious mining, no recent grazing (within 30 years), minimal timber harvest (<5%), and minimal road density (<0.5 km/km^2) (table A1).

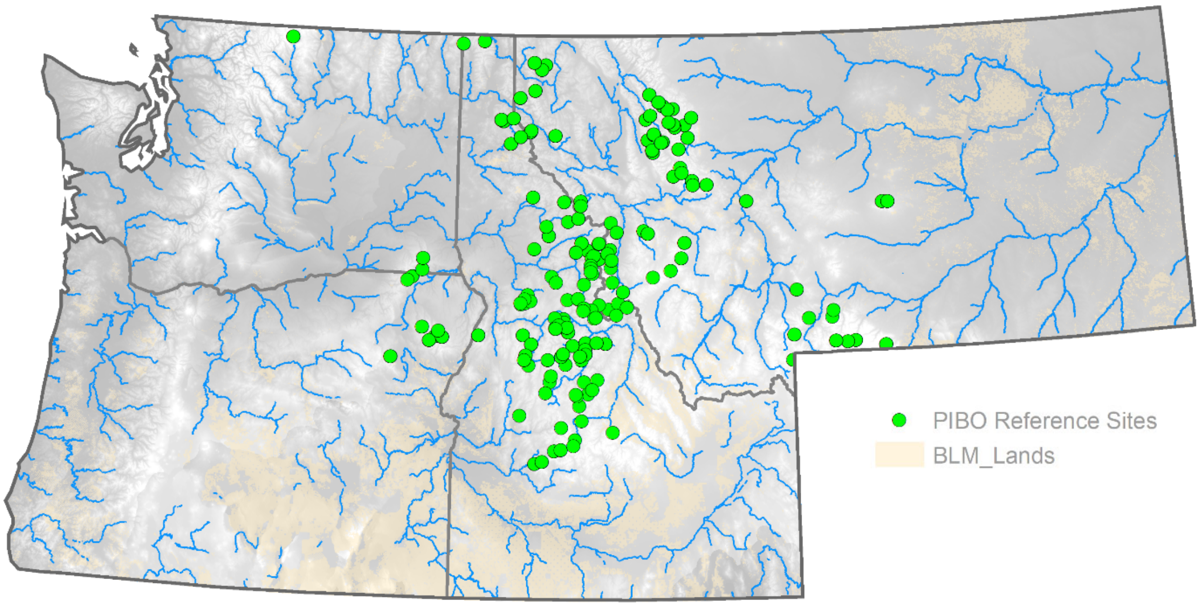


Figure A1. Spatial distribution of the 201 reference reaches used to develop the PIBO macroinvertebrate O/E model for assessing the biological condition of wadeable stream systems.

Overview of specific conductance, total nitrogen, and total phosphorus predictive models and reference criteria

For the water quality indicators lacking state standards, we used predictive models to establish PNC in the absence of anthropogenic impacts. PNC uses empirical models based on geospatial predictors to understand the spatial variability among reference reaches for a given indicator. Such models account for natural environmental gradients and predict chemical, physical, or biological values expected at a reach in the absence of anthropogenic impairment (i.e., PNC). Condition is then determined based on the deviation of the observed indicator value from the reach-specific predicted value. If this deviation is greater than specified percentiles of model error (e.g., 95th), the value is assigned a condition of "not meeting" a given water quality benchmark. Predictive modeling approaches are advantageous because they result in reach-specific predictions, take into account natural environmental gradients, and have known levels of accuracy and precision.

We established specific conductance benchmarks using the methods of Olson et al. (2012). This model uses 15 geographic information system (GIS)-derived variables (e.g., percent calcium carbonate in local geology, air temperature, precipitation) to explain 71% of the spatial variability in base-flow specific conductance concentrations (root-mean-square error 84.2 µS/cm) among reference reaches throughout the contiguous western United States. We then established reach-specific benchmarks by taking the reach-specific PNC from the model and adding the 95% of model error (74.5 mS/cm) to the prediction.

We established total nitrogen and total phosphorus benchmarks using the methods in Olson and Hawkins (2013). The total nitrogen model uses 12 GIS-derived variables (e.g., atmospheric nitrogen deposition, air temperature, precipitation) to explain 23% of the spatial variability in base-flow total nitrogen values (root-mean-square error 80.1 µg/L). The total phosphorus model uses 15 GIS-derived variables (e.g., percent calcium carbonate in local geology, air temperature, precipitation) to explain 46% of the spatial variability in base-flow total phosphorus values (root-mean-square error 20.5 µg/L) among reference reaches throughout the contiguous western United States. We established benchmarks for total nitrogen and total phosphorus similar to specific conductance. The 95% of model error is 114.7 µg/L for total nitrogen and 21.3 µg/L for total phosphorus.

Olson and Hawkins (2012, 2013) used a reference network derived by compiling data from state and national water quality monitoring efforts for which the original collection agency identified sampled reaches as being in reference quality (Olson 2012, table 2-4). Olson and Hawkins confirmed the quality of these reaches following a two-step process. First, they used field-based physical habitat and water quality data for the sampled reaches to screen data for anomalous water quality values. Second, they used Google Earth and USGS quad maps to screen reaches for any evidence of human impacts (e.g., ranches, mines, agriculture, clearcuts).

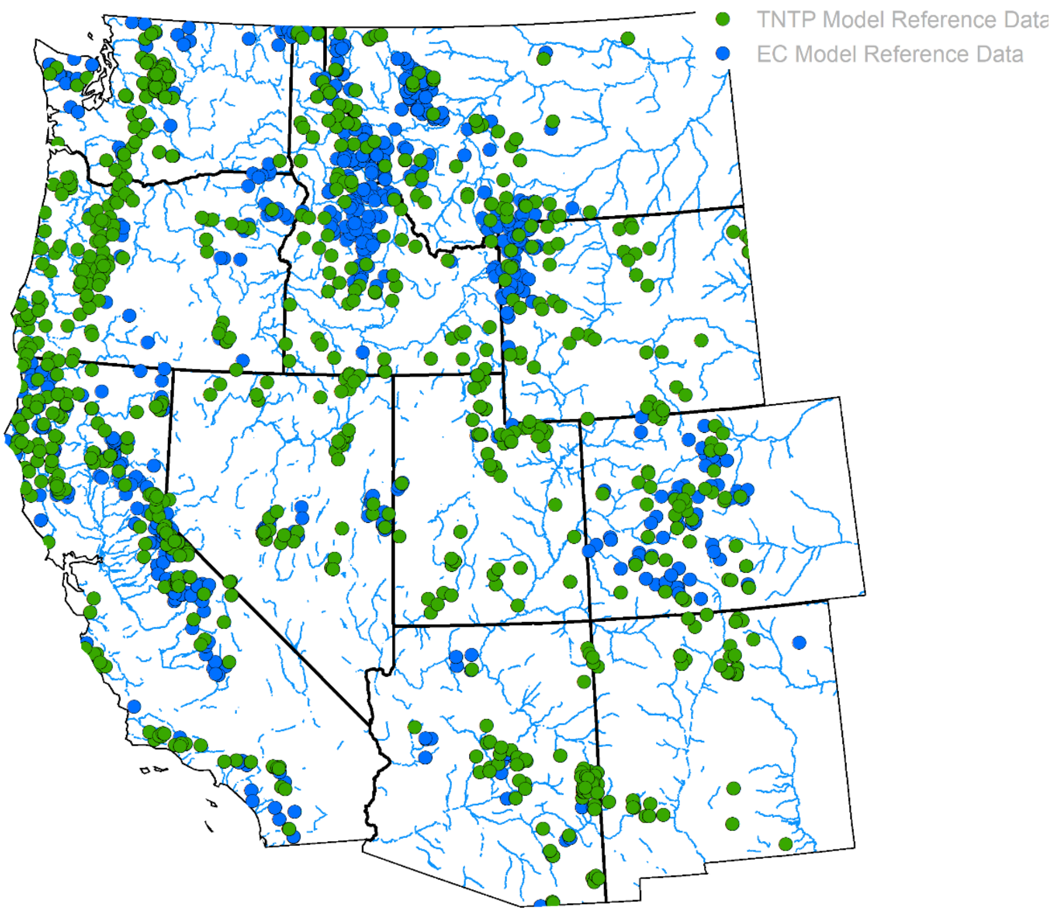


Figure A2. Spatial distribution of the reference reaches used to develop predictive models for total nitrogen, total phosphorus, and specific conductance (Olson and Hawkins 2012, 2013).

Overview of methods used to develop benchmarks for percent fine sediment, bank canopy cover, and channel incision

All instream habitat and riparian indicators generally lacked predictive models or state standards except for bank stability and cover, and therefore, we based benchmarks on the percentiles of regional reference conditions (Hughes et al. 1986; Stoddard et al. 2006). Specifically, we used EPA data from 226 reference reaches for the two hybrid level III ecoregions encompassing the state of Idaho: 185 reference reaches for the Northern Rockies hybrid ecoregion and 41 reaches for the Northern Xeric Basins (figure A3).

We used reference reaches to characterize the natural range of indicator variability expected to occur in the absence of anthropogenic impairment. We established benchmarks at the extremes of reference reach distributions to identify significant departures from reference for each of two stream sizes: small wadeable reaches (≤10 m bankfull width) and large wadeable reaches (>10 m bankfull width). For example, we used the 70th percentile of reference reach percent fine sediment values (<2 mm) for small wadeable streams in the Eastern Xeric Basins ecoregion (45%) to determine if individual reaches were meeting or not meeting the benchmark. In other words, we categorized reaches as not meeting the fine sediment benchmark if measurements exceeded levels observed among 70% (44% fine sediment) of reference reaches. Appendix B illustrates raw indicator values compared to benchmarks for each of the instream habitat and riparian indicators.

We used an EPA dataset to identify the range of variability among least disturbed reaches (i.e., reference reaches) by hybrid Omernick level III ecoregions. The reference dataset was comprised of 226 reaches sampled between 2000 and 2009 as part of EPA's Wadeable Streams Assessment (WSA) (Olsen and Peck 2008), Western Environmental Monitoring and Assessment Program (EMAP-West) (Stoddard, Peck, Olsen, et al. 2005), and National Rivers and Streams Assessment (NRSA) surveys (Herlihy et al. 2008). The EPA screening process for reference reach designations differed slightly among surveys, but we used all three datasets to maximize sample sizes within each ecoregion.

In general, the EPA used multiple lines of evidence to screen sampled reaches to determine those that were least disturbed. At the broadest scale, GIS-derived metrics of land use (e.g., row crop and urban land use) and other anthropogenic activities (e.g., dams and impoundments) were used to screen reaches (table A1). At the reach-scale, we used field observations of the magnitude and proximity of streamside human activities such as roads, agricultural and urban development, and riparian disturbance as described by Kaufmann et al. (1999) and Herlihy et al. (2008). We also used instream habitat variables such as bank canopy cover and fine sediment levels. Lastly, we used GIS and field-based observations in conjunction with water chemistry data, as described by Herlihy et al. (2008), to designate reaches as least, moderately, and highly disturbed relative to other sampled reaches.

To avoid circularity in the above process, we did not use any measures directly related to a given indicator to screen reaches. For example, we did not use field measurements of riparian vegetation, sediment, or instream habitat complexity to screen reaches and determine ranges of variability for any instream indicators. The only exceptions were the use of riparian conditions for designations of instream habitat indicators and vice versa.

Table A1. Reference reach screening criteria used by the three different networks.

Reference screening criteria	PIBO	EPA	Olson and Hawkins[1]
Grazing	none in last 30 years	NA	Visual assessment using aerial imagery—no numeric criteria applied
Timber harvest	<5%	NA	
Road density	<0.5 km/km^2	<1 km/km^2	
Percent ag land	NA, but generally 0%	<3%	<5%
Percent urban	NA, but generally 0%	<3%	<5%
Percent ag + urban	NA, but generally 0%	<5%	NA
Dam density	NA, but generally 0%	<0.005	Visual assessment using aerial imagery—no numeric criteria applied
Mine density	NA, but generally 0%	0	
National Pollutant Discharge Elimination System	NA, but generally 0%	0	
Riparian human disturbance index	NA, but generally 0%	<1.5	<2

[1]Olson and Hawkins (2012, 2013) also used field-based water quality criteria for establishing reference conditions following the guidance of Herlihy et al. (2008).

Given a lack of predictive models of instream habitat and riparian indicators, BLM's national AIM team attempted to minimize natural variability associated with reaches within a given hybrid ecoregion. Similarly to the approach taken by the EPA in the Western Environmental Monitoring and Assessment Program (EMAP-West) surveys, we used EPA hybrid level II/III ecoregions to divide reference reaches into relatively homogenous physiographic regions. Then within a given ecoregion, we used bankfull width to separate reference reaches into small streams (≤10 m bankfull width) and large streams (>10 m). We chose 10 m as an arbitrary cutoff based on balancing sample sizes and maximizing discriminatory efficiency for individual indicators between the two groups.

Using this approach, most indicators had a substantial difference between benchmark values for small streams and large streams that made ecological sense while still providing adequate sample sizes for most indicators for a given ecoregion and stream size (e.g., >20 reaches). For example, the benchmark for major departure from reference for percent canopy cover in the Northern Rockies is 20.9% for small streams (which generally support more canopy cover than large streams) but 2.3% for large streams.

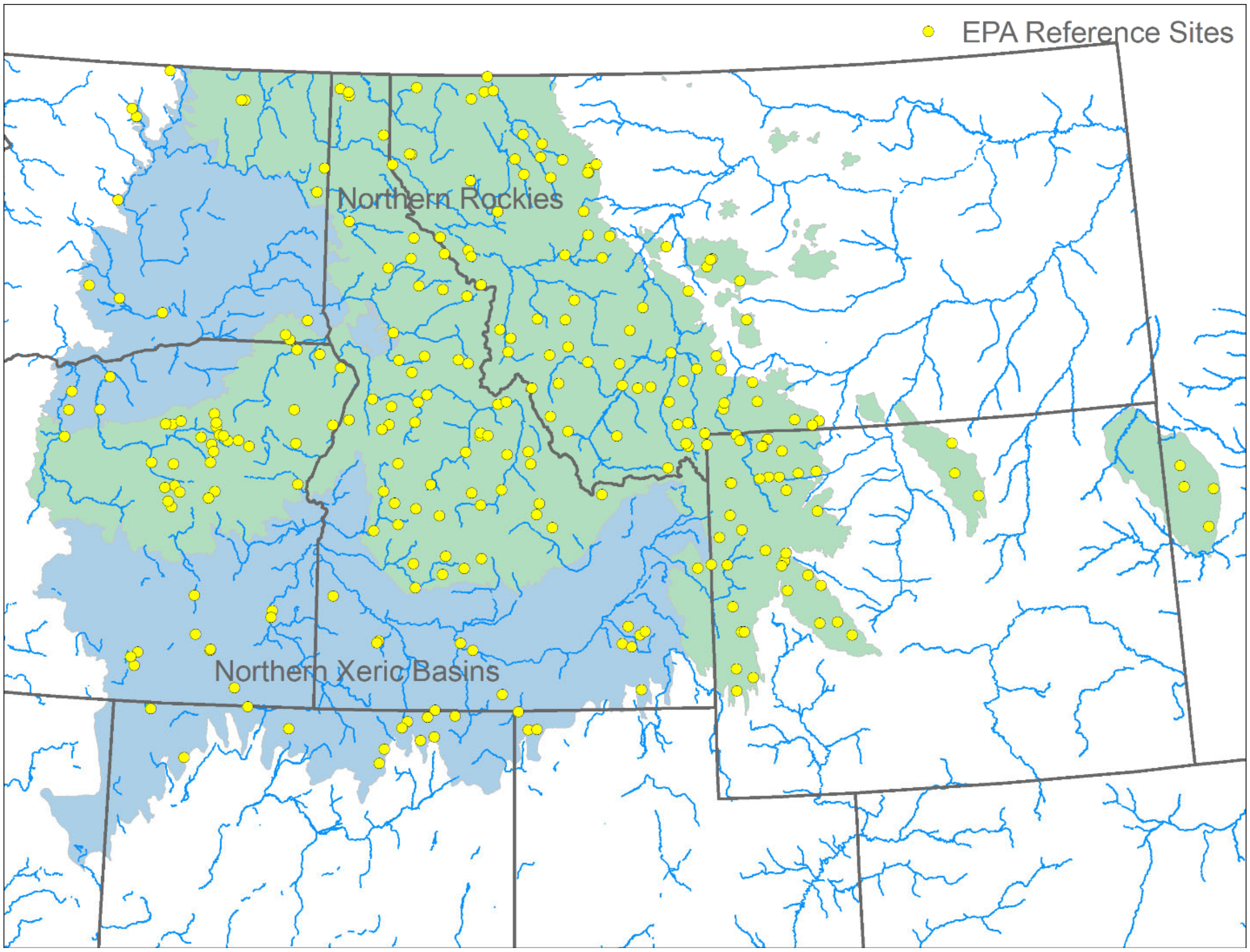
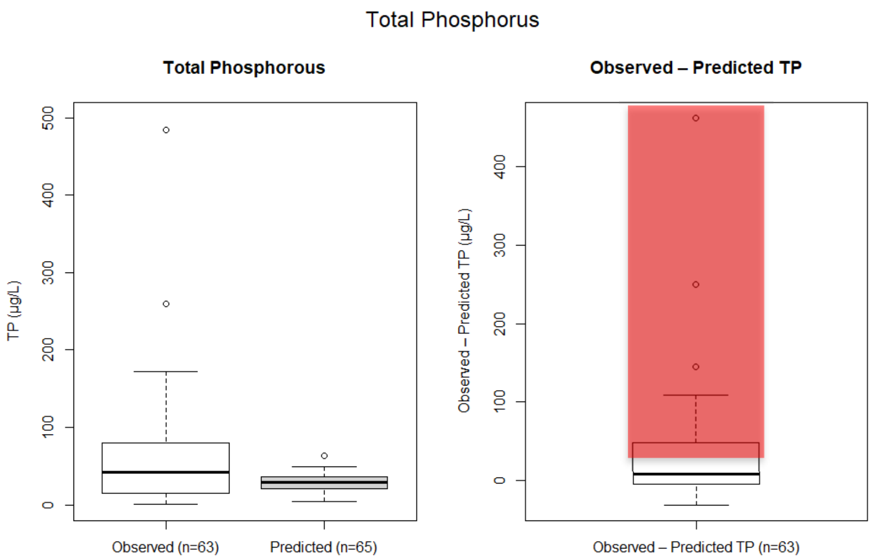
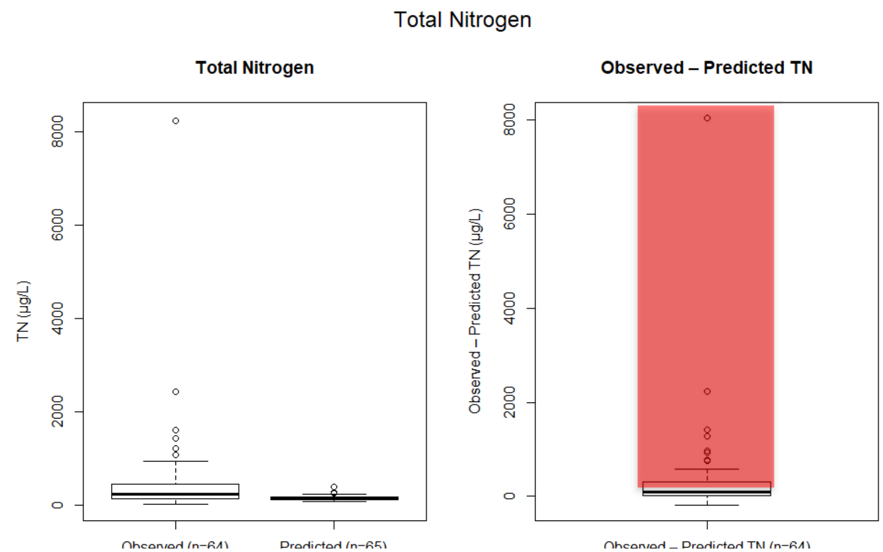


Figure A3. Spatial distribution of the 226 reference reaches used to characterize the natural range of variability among the Northern Rockies (n = 185) and Northern Xeric Basins (n = 41) hybrid ecoregions for percent fine sediment, bank canopy cover, and channel incision.

Appendix B: Box Plots of Raw Indicator Values Compared to Benchmarks





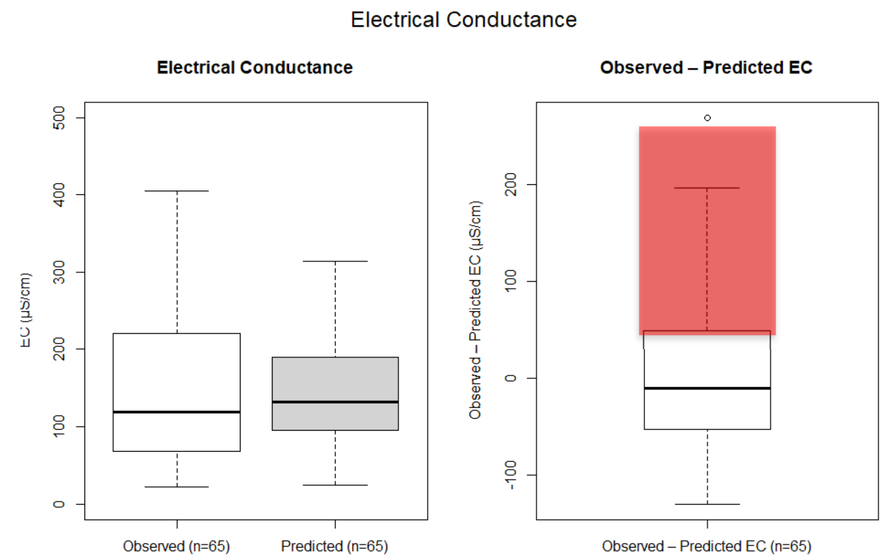


Figure B1. Total phosphorus, total nitrogen, and specific conductance values compared between observed and predicted conditions across the 65 sampled reaches as well as the difference between observed and predicted conditions, with benchmark exceedances for test reaches highlighted in red. Box plots are a standardized representation of data based on five descriptive statistics: minimum values (bottom whisker shown as 1.5 times the inner quartile range), first quartile (bottom of box), median (bold line in box), third quartile (top of box), and maximum value (top whisker shown as 1.5 times the inner quartile range). Individual circles represent outlier values.

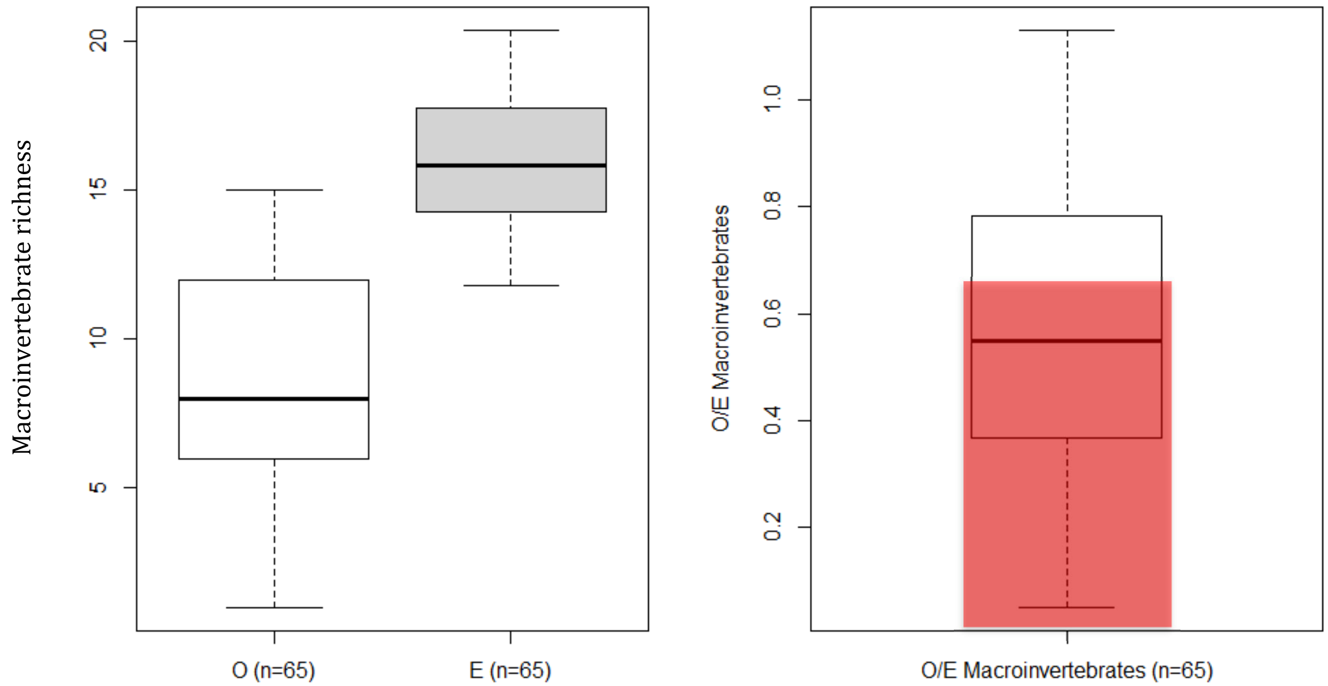


Figure B2. Macroinvertebrate biological condition values compared between observed (O) and predicted (E) conditions across the 65 sampled reaches as well as the ratio of observed to predicted conditions, with benchmark exceedances for test reaches highlighted in red. See figure B1 for a general box plot description.

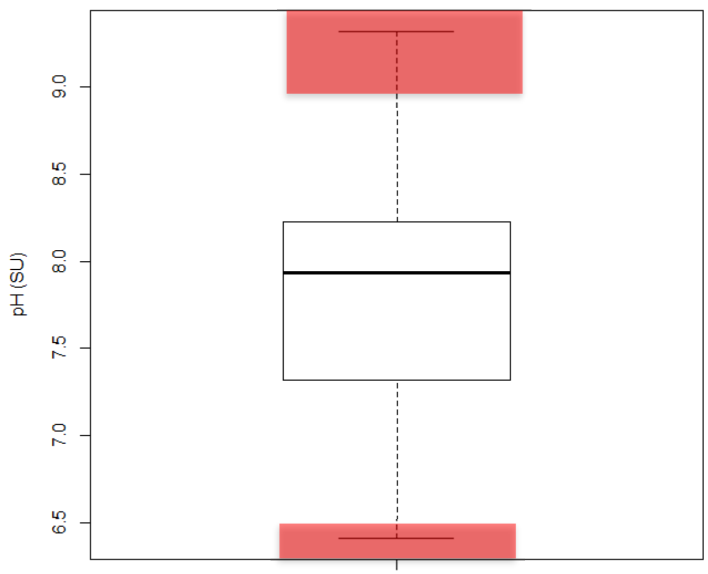


Figure B3. The pH values observed among the 65 sampled BLM-managed stream reaches throughout Idaho. The upper (alkaline) and lower (acidic) red boxes indicate benchmark exceedances for test reaches per Idaho DEQ guidance. See figure B1 for a general box plot description.

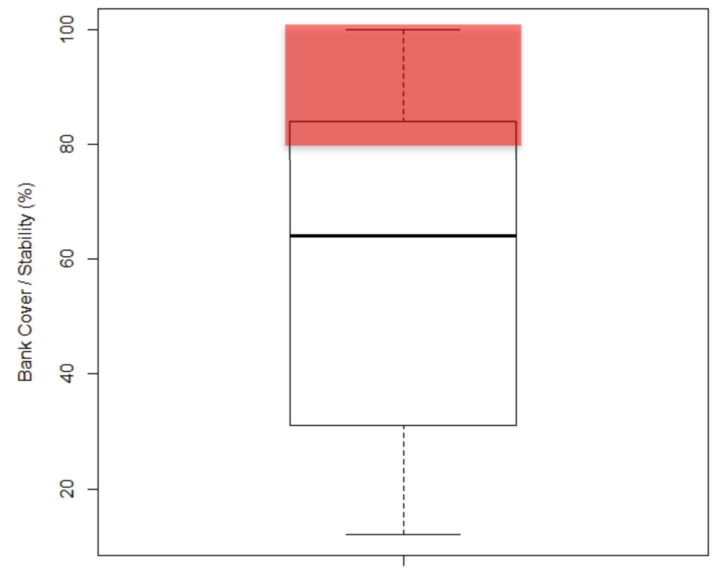


Figure B4. Combined bank cover/stability values observed among the 65 sampled BLM-managed stream reaches throughout Idaho. The red box indicates benchmark exceedances for test reaches (≥80% of plots at a reach categorized as unstable or uncovered). See figure B1 for a general box plot description.

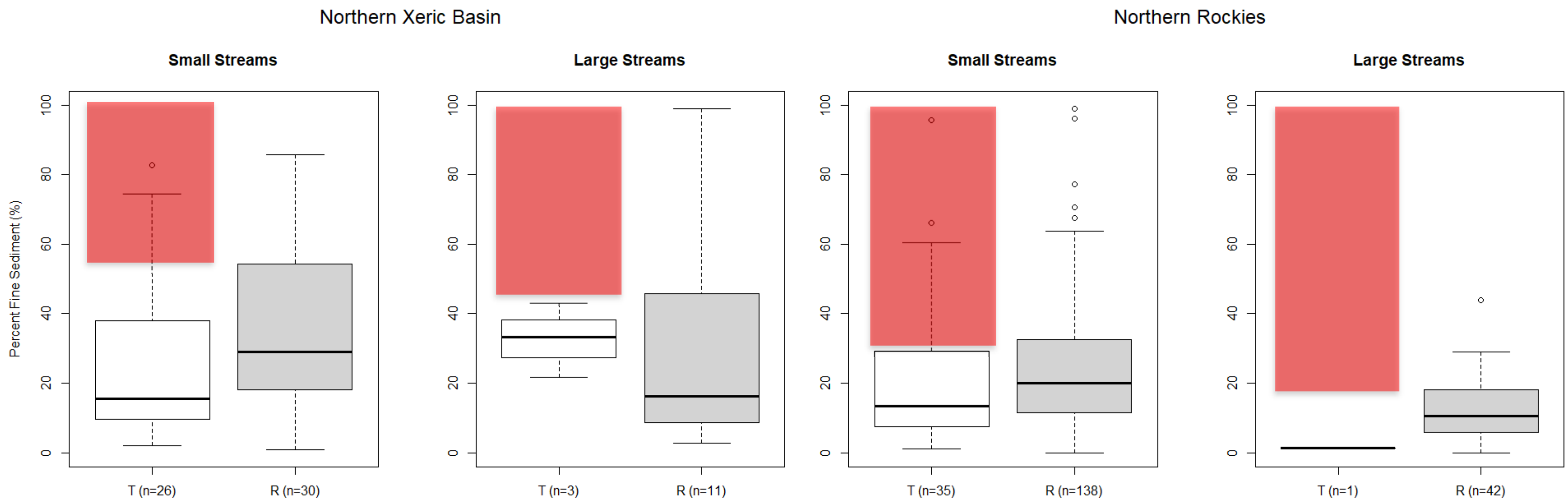


Figure B5. Percent fine sediment compared between reference (R) and sampled test reaches (T) reaches for both the Northern Xeric Basins and Northern Rockies hybrid ecoregions for small (≤10 m) and large (>10 m) streams. The red box indicates benchmark exceedances for sampled test reaches (≤10th percentile of reference distributions). Note that the ecoregion and stream size groups are not reporting units but rather groups developed for setting differential benchmarks among systems based on potential. See figure B1 for a general box plot description.

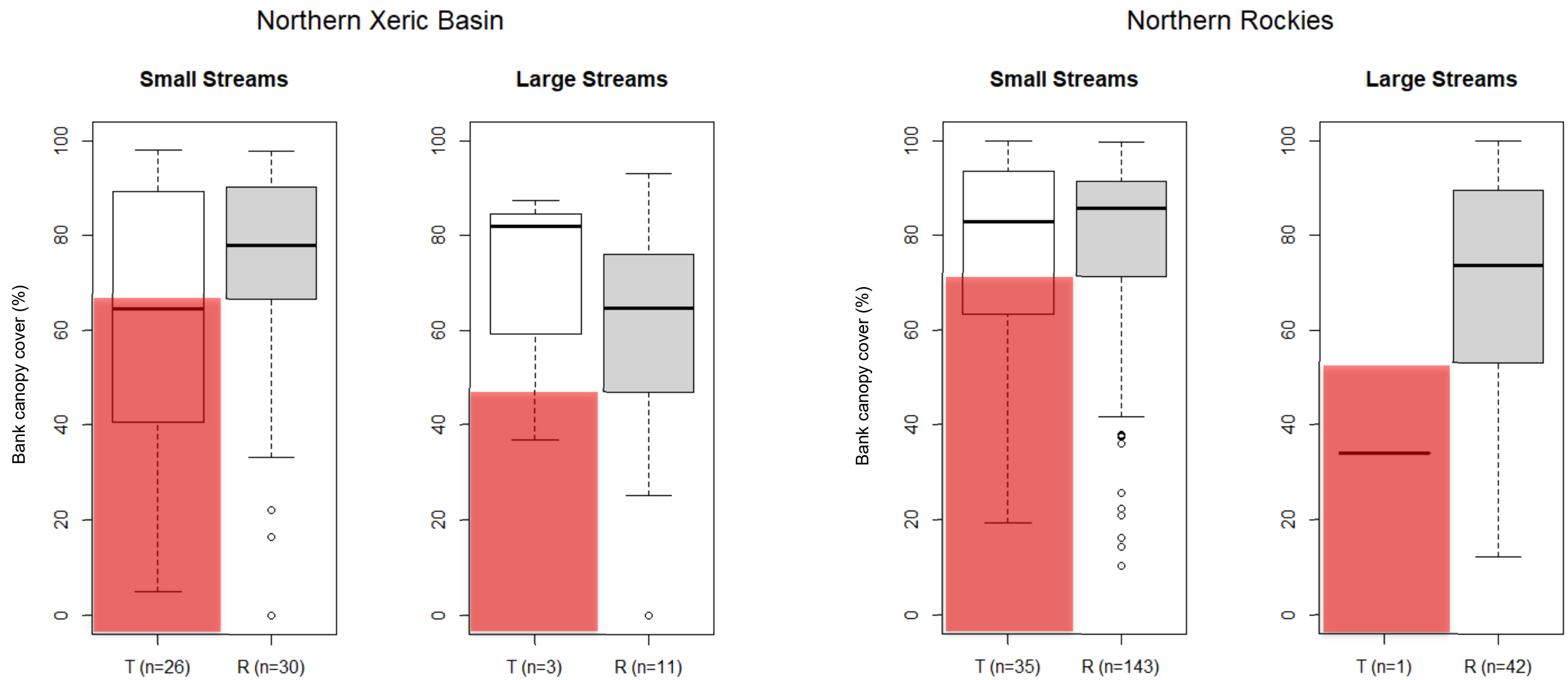


Figure B6. Percent bank canopy cover compared between reference (R) and sampled test (T) reaches for both the Northern Xeric Basins and Northern Rockies hybrid ecoregions and small (≤10 m) and large (>10 m) streams. The red box indicates benchmark exceedances for sampled test reaches (≤10th percentile of reference distributions). Note that the ecoregion and stream size groups are not reporting units, but rather groups developed for setting differential benchmarks among systems based on potential. See figure B1 for a general box plot description.

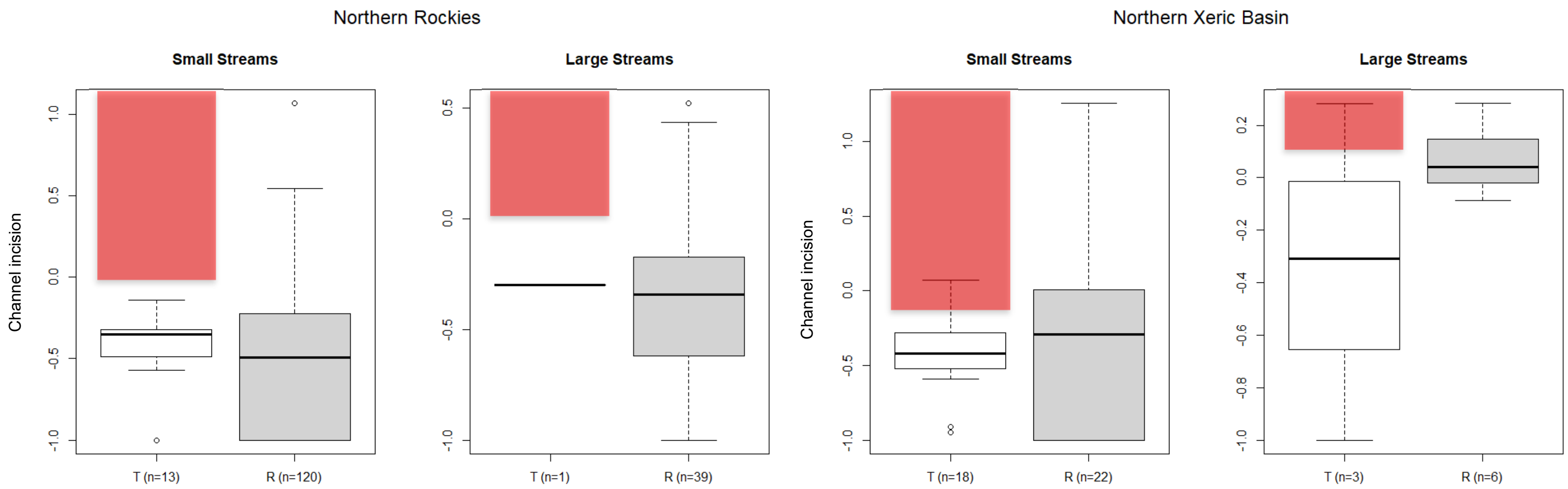


Figure B7. Channel incision compared between reference (R) and sampled test (T) reaches for both the Northern Xeric Basins and Northern Rockies hybrid ecoregions and small (≤10 m) and large (>10 m) streams. Benchmark exceedances for sampled test reaches are indicated by the red box (≤10th percentile of reference distributions). Note that the ecoregion and stream size groups are not reporting units, but rather groups developed for setting differential benchmarks among systems based on potential See figure B1 for a general box plot description.

Literature Cited

Allan, J.D. and M.M. Castillo. 2007. Stream ecology: Structure and function of running waters. Second edition. Springer.

Beschta, R.L. 1997. Riparian shade and stream temperature: An alternative perspective. Rangelands 19(2):25–28.

Bjornn, T.C. and D.W. Reiser. 1991. Habitat requirements of salmonids in streams. In *Influences of Forest and Rangeland Management on Salmonid Fishes and Their* Habitats. Edited by W.R. Mehan. American Fisheries Society Special Publication 19:83–138.

Booth, D.B. 1990. Stream-channel incision following drainage-basin urbanization. Water Resources Bulletin 26(3):407–417.

Bureau of Land Management (BLM). 2015. AIM national aquatic monitoring framework: Introducing the framework and indicators for lotic systems. Technical Reference 1735-1. U.S. Department of the Interior, Bureau of Land Management, National Operations Center, Denver, CO.

Bureau of Land Management (BLM). 2017. AIM national aquatic monitoring framework: Field protocol for wadeable lotic systems. Technical Reference 1735-2. U.S. Department of the Interior, Bureau of Land Management, National Operations Center, Denver, CO.

Cummins, K.W. 1974. Structure and function of stream ecosystems. BioScience 24(11):631–641.

Cunjak, R. A. and G. Power. 1986. Winter habitat utilization by stream resident brook trout (*Salvelinus fontinalis*) and brown trout (*Salmo trutta*). Canadian Journal of Fisheries and Aquatic Sciences 43(10):1970–1981.

Dodds, W.K. 2006. Eutrophication and trophic state in rivers and streams. Limnology and Oceanography 51:671–680.

Dunck, B., E. Lima-Fernandes, F. Cássio, A. Cunha, L. Rodrigues, and C. Pascoal. 2015. Responses of primary production, leaf litter decomposition and associated communities to stream eutrophication. Environmental Pollution 202:32–40.

Hawkins, C.P., R.H. Norris, J.N. Hogue, and J.W. Feminella. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecological Applications 10(5):1456–1477.

Henley, W.F., M. . Patterson, R.J. Neves, and A.D. Lemly. 2000. Effects of sedimentation and turbidity on lotic food webs: A concise review for natural resource managers. Reviews in Fisheries Science 8(2):125–139.

Herbst, D.B., M.T. Bogan, S.K. Roll, and H.D. Safford. 2012. Effects of livestock exclusion on in-stream habitat and benthic invertebrate assemblages in montane streams. Freshwater Biology 57(1):204–217.

Herlihy, A.T., S.G. Paulsen, J. Van Sickle, J.L. Stoddard, C.P. Hawkins, and L.L. Yuan. 2008. Striving for consistency in a national assessment: The challenges of applying a reference-condition approach at a continental scale. Journal of the North American Benthological Society 27(4):860–877.

Hogsden, K.L. and J.S. Harding. 2012. Consequences of acid mine drainage for the structure and function of benthic stream communities: A review. Freshwater Science 31(1):108–120.

Hughes, R.M., D.P. Larsen, and J.M. Omernik. 1986. Regional reference reaches: A method for assessing stream potentials. Environmental Management 10(5):629–635.

Isaak, D.J., S.J. Wenger, E.E. Peterson, J.M. Ver Hoef, D.E. Nagel, C.H. Luce, S.W. Hostetler, J.B. Dunham, B.B. Roper, S.P. Wollrab, G.W. Chandler, D.L. Horan, and S. Parkes-Payne. 2017. The NorWeST summer stream temperature model and scenarios for the western U.S.: A crowd-sourced database and new geospatial tools foster a user community and predict broad climate warming of rivers and streams. Water Resources Research 53(11):9181–9205.

Irvine, K, S.W. Miller, R. Al-Chokhachy, B. Roper, J. Kershner. 2015. Empirical evaluation of the conceptual underpinnning a regional aquatic long-term monitoring program using causal modelling. Ecological Indicators 50: 8-23.

Johnson, S.L. and J.A. Jones. 2000. Stream temperature responses to forest harvest and debris flows in western Cascades, Oregon. Canadian Journal of Fisheries and Aquatic Sciences 57(S2):30–39.

Kaufmann, P.R., P. Levine, E.G. Robison, C. Seeliger, and D.V Peck. 1999. Quantifying physical habitat in wadeable streams. EPA/620/R-99/003. U.S. Environmental Protection Agency, Washington, DC.

Knapp, R.A. and K.R. Matthews. 1996. Livestock grazing, golden trout, and streams in the Golden Trout Wilderness, California: Impacts and management implications. North American Journal of Fisheries Management 16(4):805–820.

Lyons, J., B.M. Weigel, L.K. Paine, and D.J. Undersander. 2000. Influence of intensive rotational grazing on bank erosion, fish habitat quality, and fish communities in southwestern Wisconsin trout streams. Journal of Soil and Water Conservation 55(3):271–276.

Miller, S.A. and T.A. Crowl. 2006. Effects of common carp (*Cyprinus carpio*) on macrophytes and invertebrate communities in a shallow lake. Freshwater Biology 51(1):85–94.

Olsen, A.R. and D.V. Peck. 2008. Survey design and extent estimates for the Wadeable Streams Assessment. Journal of the North American Benthological Society 27(4):822–836.

Olson, J.R. 2012. The influence of geology and other environmental factors on stream water chemistry and benthic invertebrate assemblages. Utah State University, Logan, UT.

Olson, J.R. and C.P. Hawkins. 2012. Predicting natural base-flow stream water chemistry in the western United States. Water Resources Research 48(2):W02504.

Olson, J.R. and C.P. Hawkins. 2013. Developing reach-specific nutrient criteria from empirical models. Freshwater Science 32(3):719–740.

Ripple, W.J. and R.L. Beschta. 2006. Linking a cougar decline, trophic cascade, and catastrophic regime shift in Zion National Park. Biological Conservation 133(4):397–408.

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. U.S. 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 condition. Ecological Applications 16(4):1267–76.

Strahler, A. N. 1952. Hypsometric (area-altitude) analysis of erosional topograph. Bulletin of the Geological Society of America 63:1117–1142.

Tank, S.E., L.F.W. Lesack, and D.J. McQueen. 2009. Elevated pH regulates bacterial carbon cycling in lakes with high photosynthetic activity. Ecology 90(7):1910–1922.

U.S. DOI BLM. 1997. Idaho standards for rangeland health and guidelines for livestock grazing

management. BLM/ID/PT-97/002+4120 rev 8/97.
Van Sickle, J., J.L. Stoddard, S.G. Paulsen, and A.R. Olsen. 2006. Using relative risk to compare the effects of aquatic stressors at a regional scale. Environmental Management 38(6):1020–1030.
Wood, P.J. and P.D. Armitage. 1997. Biological effects of fine sediment in the lotic environment. Environmental Management 21(2):203–217.
Zaimes, G.N., R.C. Schultz, and T.M. Isenhart. 2004. Stream bank erosion adjacent to riparian forest buffers, row-crop fields, and continuously-grazed pastures along Bear Creek in central Iowa. Journal of Soil and Water Conservation 59(1):19–27.