Big Hole River Arctic Grayling Risk Assessment

Upper Missouri River Distinct Population Segment (Thymallus arcticus)

Prepared for: Center for Biological Diversity

Prepared by: ECOREGIONS LLC

Authors: Gary Rule, Richard Domingue, and Blane Bellerud

Date: January 21, 2026

Author credentials

Gary Rule: Fisheries biologist with NOAA Fisheries (2001-2025) specializing in ESA status reviews, critical habitat analyses, and scientific permitting; Master of Environmental Management (Portland State University).

Richard A. Domingue: Hydrologist and instream-flow specialist; Bachelor of Science in Forest and Watershed Management (Colorado State University, 1979); former National Marine Fisheries Service hydrologist (1998-2017) supporting ESA consultations for major dam and hydropower systems.

Blane Bellerud: Fisheries biologist with experience at Oregon Department of Fish and Wildlife (6 years) and NOAA Fisheries (23 years); M.S. Biology and Ph.D. in Aquatic Parasitology; author of peer-reviewed publications.

Table of Contents

List of Acronyms

Acronym	Definition
A_r	Allelic richness
CBD	Center for Biological Diversity
CCAA	Candidate Conservation Agreement with Assurances
CI	Confidence interval
CPUE	Catch per unit effort
CTM	Critical thermal maximum
DNRC	Montana Department of Natural Resources and Conservation
DPS	Distinct Population Segment
ESA	Endangered Species Act
GIS	Geographic information system
H_e	Expected heterozygosity
MFWP	Montana Fish, Wildlife & Parks
MK	Mann–Kendall (trend test)
NAD83	North American Datum of 1983
N_b	Effective number of breeders
N_e	Effective population size
OR	Odds ratio
UILT	Upper incipient lethal temperature
USFWS	U.S. Fish and Wildlife Service
USGS	U.S. Geological Survey
WGS84	World Geodetic System 1984
YOY	Young-of-year

List of Tables

List of Figures

Executive Summary

Based on the evidence reviewed in this report, we conclude that the Upper Missouri River Arctic grayling Distinct Population Segment (DPS) warrants listing under the Endangered Species Act (ESA). Viability has not improved across multiple independent indicators. Key threats—especially chronic summer/fall (July–October) low flows and elevated stream temperatures—remain persistent and, in some cases, are worsening.

Recent genetic syntheses describe the DPS as a mix of native and introduced populations, including the remaining fluvial population in the upper Big Hole River basin and multiple introduced mountain-lake populations in the basin. This report does not evaluate the conservation value of introduced lake populations; it focuses on the Big Hole River because it retains the DPS's strongest remaining fluvial expression. Loss of this population would materially reduce the DPS's adaptive capacity.

We evaluated multiple, independent indicators of population status and habitat condition, including effective number of breeders (N_b), catch per unit effort (CPUE), genetic diversity metrics, and long-term performance against seasonal instream-flow and stream-temperature benchmarks. Because USFWS has relied on genetic effective population size (N_e) as supporting evidence in prior findings, we briefly interpret effective population size alongside these indicators, rather than as a substitute for demographic trends and threat abatement.

Key Findings

Demographic indicators (N_b and CPUE):

- N_b estimates indicate, at most, weak or episodic improvement and do not show a sustained upward trend consistent with recovery.
- Fall CPUE has not shown a consistent increase in abundance of Age-1+ fish
- Periodic young-of-year (YOY) recruitment pulses have not translated into persistent gains in older fish.

Genetic diversity:

- The frequently cited "<10% heterozygosity loss" framing is not, by itself, a reliable proxy for demographic security.
- Expected heterozygosity can remain relatively stable even as allelic richness declines following bottlenecks and genetic drift.
- Recent monitoring indicates the effective population size (N_e) for the Big Hole River is below widely cited long-term benchmarks for maintaining adaptive potential (i.e., retaining allelic richness and adaptive capacity).
- Accordingly, recent N_e estimates do not support a conclusion of long-term genetic security for the Big Hole River population.

Instream flow:

- Seasonal instream-flow targets under the Big Hole River Candidate Conservation Agreement with Assurances (CCAA) are not met with sufficient frequency, and several segments show declining compliance over time.
- Chronic summer/fall shortfalls are especially consequential because they coincide with the warmest period and highest biological stress.

Stream temperature:

- Recurrent summer/fall exceedances of the chronic stress threshold (e.g. 21°C) potentially reduce usable habitat and increase physiological stress.

Synthesis

In our professional judgment, the weight of evidence supports listing the Upper Missouri River DPS under the Endangered Species Act. Across demographic indicators, flow, temperature, and genetics, the record indicates continued vulnerability to the same core threats that have threatened the species survival since at least the 1980’s. The Big Hole River population remains a critical component of DPS representation and adaptive capacity.

Background and Purpose

For decades, federal status reviews and basin conservation planning have treated the Big Hole River Arctic grayling population as a high-risk remnant: a small population confined to a tiny fraction of its historic distribution and especially vulnerable during late summer. Concern has consistently centered on two linked stressors that directly limit usable river habitat during the warm season—reduced instream flows and elevated stream temperatures. Accordingly, this report asks whether these two core stressors (flow and temperature) have improved and whether the grayling population shows any corresponding evidence of recovery.

ESA Listing History and Litigation Context

Upper Missouri River Arctic grayling have a long ESA history. In 1994, the U.S. Fish and Wildlife Service (USFWS) first concluded that listing the Upper Missouri River fluvial Arctic grayling was warranted, and later reached the same conclusion again following renewed review (USFWS 1994 and 2010). Between those actions, the Service also elevated the listing priority number to reflect imminent, high-magnitude threats (USFWS 2004).

Since the 2010 "warranted" determination, USFWS has made two "not warranted" determinations which were rejected in federal court. The 2014 finding was vacated by the Ninth Circuit, which identified multiple errors, including failure to adequately assess the Big Hole River population decline and an arbitrary determination regarding Ruby River viability. After USFWS again concluded listing was not warranted in 2020, the U.S. District Court for the District of Montana vacated the 2020 finding and remanded for a new determination. The court held the 2020 finding was arbitrary and capricious insofar as it treated the Ruby River population as viable and capable of providing redundancy. The court also faulted USFWS for relying on the benefits of the Big Hole CCAA without considering whether the CCAA (and at least some benefits) could cease to exist in the foreseeable future or whether site-specific plans would be renewed.

In our view, little has changed in the fundamental risk profile during the CCAA era. The threats most strongly linked to demographic performance—summer/fall low flows and elevated summer temperatures—remain active at the basin scale, and available monitoring does not demonstrate a measurable improvement in overall viability.

Historical Decline and Causes

The modern decline was already evident in field monitoring by the 1980s: population estimates in the upper Big Hole fell sharply from early-1980s highs to late-1980s lows. One long-running monitoring record summarizes a decline from 111 fish per mile (1983) to 22 fish per mile (1989) and notes that the population appeared to stabilize only after several years of decline (Byorth 1991). Investigators working during that period repeatedly flagged dewatering and the severe 1988–1989 drought as likely contributors (alongside other pressures) (Streu 1990). In response to recurring low-flow and high-temperature conditions, local partners adopted a basin-wide drought response framework in the 1990s intended to mitigate low stream flows and lethal water temperatures for fisheries—"particularly fluvial Arctic grayling"—including flow-based triggers tied to voluntary reductions in irrigation and temperature-based triggers keyed to prolonged exposure above 70°F (Big Hole River Watershed Committee 1997).

Scope and Population Focus

Recent syntheses and genetic analyses provide a clear description of DPS population structure. The Montana Arctic Grayling Workgroup (2022) and Kreiner (2025) identify 19 extant populations across the broader conservation portfolio. These include historically occurring, naturally established populations in the upper Big Hole and Centennial valleys, and 12 introduced mountain-lake populations that are genetically isolated from other populations.

A central question is whether introduced mountain-lake populations should be credited in DPS viability evaluations. USFWS's own approach has varied over time—from excluding introduced populations to later including lake populations in the DPS and citing their presumed climate resilience in "not warranted" findings.

Our report does not try to resolve how introduced lake populations should be weighted in conservation decisions; instead, it focuses on the Big Hole River because it is central to the DPS's remaining fluvial life-history expression and because it is a focal population in prior petitioning and agency analyses. The 2006 Big Hole River Candidate Conservation Agreement with Assurances (CCAA) describes grayling as once widespread across the upper Missouri drainage but now occupying less than 5% of their former range, with the Big Hole population largely concentrated in a core portion of the river (USFWS 2006).

The chapters that follow evaluate viability in the Big Hole River population and the key threats directly tied to persistence (streamflow and temperature). Where relevant, we also reference N_e because it has been central to USFWS's genetics-based extinction risk reasoning; we treat it as supporting genetic context rather than as a stand-alone extinction risk indicator.

This assessment was prepared for the Center for Biological Diversity (CBD) to support its ongoing efforts concerning ESA listing of Upper Missouri River Arctic grayling.

Viability of the Big Hole River Population

In this section we review trends in abundance and genetic diversity.

Demographics

To begin, we ask a practical question: does the available information indicate improvement in the Big Hole River Arctic grayling population's viability? To answer this question, we use two independent abundance indicators collected by Montana Fish, Wildlife & Parks: (1) N_b, a genetic index tied to successful breeding and early recruitment, and (2) CPUE, a field-based index of relative abundance from standardized electrofishing surveys. We first summarize how each metric is generated and what it can—and cannot—tell us, then test for long- and short-term trends, and finally explain why N_b and CPUE can move in different directions. Across both indicators, we found that there is no clear evidence of sustained improvement in overall abundance over the monitoring period.

Effective Number of Breeders

The N_b values analyzed in this report come from a long-running genetic monitoring program led by Montana Fish, Wildlife & Parks and the University of Montana. These data are derived from annual genetic samples of Arctic grayling collected in the Big Hole River basin and are intended to track population status and trend over time. The effective number of breeders (N_b) represents the number of adult grayling that successfully contribute offspring to a given year's cohort. In plain language, N_b reflects how many fish are effectively reproducing, rather than how many fish are present in the river. Not all adults reproduce successfully every year, therefore N_b is typically smaller than the total number of spawning adults and varies from year to year.

Effective population size (N_e) is closely related but operates over longer time scales. Whereas N_b reflects genetic contributions within a single reproductive season, N_e represents the effective size of the population across generations and is more directly tied to long-term genetic health. N_e includes the retention of genetic diversity and resistance to inbreeding. In the Big Hole River, N_e estimates are derived from the long-term pattern of N_b values and species-specific life history information (Kovach 2019). As such, N_e should be interpreted as a summary indicator of long-term population resilience rather than a direct estimate of adult abundance.

It is important to recognize that N_b and N_e are genetic indicators, not direct counts of fish. Year-to-year changes in N_b can reflect changes in the number of breeders, changes in reproductive success, or both. High or low N_b values in individual years do not necessarily indicate corresponding changes in adult abundance, and short-term fluctuations are expected even in relatively stable populations. In addition, precision varies among years depending on sample size, meaning that trends are best evaluated across multiple years rather than from single-year estimates. N_b is inferred from offspring in a given year, so it is best interpreted as an index of reproductive contribution rather than a direct estimate of adult abundance.

Taken together, N_b and N_e are most informative when used to evaluate directional change and long-term stability, rather than as stand-alone measures of population recovery or decline. In this report, N_b is analyzed as a time series to test for long-term and short-term trends, with the goal of assessing whether the genetic contribution of breeders in the Big Hole River Arctic grayling population shows evidence of sustained increase, decrease, or stability through time. Interpretation of these results should be made in conjunction with complementary demographic, hydrologic, and habitat information, rather than as a proxy for census population size alone.

Effective population size context

Recent annual monitoring updates from Montana Fish, Wildlife & Parks and the University of Montana Conservation Genetics Laboratory estimate the Big Hole River genetic effective population size at N_e = 324.7 based on the full N_b time series through 2024 (Kovach et al. 2025). This aligns closely with the peer-reviewed estimate from Kovach et al. (2019), which reported an overall N_e of 291.6 for the Big Hole River based on the harmonic mean of multi-year N_b estimates and life-history ratios. Together, these estimates support USFWS's use of an N_e on the order of ~300 for drift projections, while underscoring that N_e is a long-term genetic context metric—not a direct measure of adult abundance or a stand-alone indicator of recovery.

Data and analysis windows

We evaluated two time windows: (1) a long-term window covering all available years (2007–2024; n=18) and (2) a short-term window intended to approximate two generations (2017–2024; n=8). The long-term window covers the monitoring period available for this analysis and falls largely within the CCAA era. The short-term window aligns with Big Hole River basin age structure: most observed spawners are age 3–4 and older fish are uncommon; therefore, ~8 years is a reasonable two-generation frame for detecting recent changes relevant to management.

These two analytical windows serve different but complementary purposes. The long-term window provides context for evaluating whether N_b has changed directionally since genetic monitoring began and since implementation of conservation measures under the Big Hole River CCAA. In contrast, the two-generation window emphasizes recent conditions that are most relevant to current management decisions and reflects the population's contemporary reproductive output. Evaluating both windows allows separation of long-term background variability from short-term signals that could indicate emerging change or stabilization.

Descriptive summary

We summarized the distribution of N_b for each window to describe central tendency and variability (Table 1). Across 2007–2024, N_b had a median of 138 and ranged from 77 to 433, indicating substantial interannual variability and occasional high years (notably 2012–2014 and 2019). In the two-generation window (2017–2024), the median was 152 and N_b ranged from 117 to 333, suggesting a higher central tendency in recent years but continued variability. The observed interannual variability in N_b is expected for a species with variable recruitment success and environmental sensitivity. High N_b years likely reflect favorable spawning and early rearing conditions, whereas low N_b years may result from reduced spawning participation, lower reproductive success, or early life-stage mortality. Importantly, variability alone does not imply instability; rather, it highlights the need for trend-based analyses that distinguish consistent directional change from short-term fluctuation.

Table 1. Summary statistics.

Statistic	All Years (2007–2024)	Last 8 Years (2017–2024)
Mean N_b	177	184
Median N_b	161	170
Minimum N_b	77	117
Maximum N_b	433	333

Methods

We tested for monotonic trends (consistent increases or decreases through time) using the two-sided Mann–Kendall (MK) test and quantified trend magnitude using Sen's slope. Because abundance time series often exhibit serial dependence (the abundance of a species in one year influences its abundance in subsequent years), we calculated lag-1 autocorrelation (r_1) and applied the Modified Mann–Kendall test (Hamed & Rao 1998), which adjusts the MK variance using an effective sample size. We treated $p < 0.05$ as statistically significant. MK asks whether later years tend to be higher (or lower) than earlier years, without assuming a normal distribution. Sen's slope estimates the typical year-to-year change. Modified MK prevents overstating significance when consecutive years are similar.

Because the MK framework evaluates the consistency of ordering through time rather than the magnitude of individual observations, it is robust to outliers and episodic high or low years. This property is particularly relevant for N_b time series, which may include occasional recruitment pulses that can strongly influence visual trends but do not necessarily represent sustained population change. The combined use of standard and modified MK tests allows explicit evaluation of whether temporal dependence affects inference.

Results

Results of the long- and short-term trend tests are summarized in Table 2 and Figures 1 and 2. Together, these results evaluate whether N_b exhibits consistent directional change over the full monitoring period and over the most recent two generations. In the figures, annual estimates of N_b are shown as points connected by lines. The dashed line represents Sen's slope; a robust estimate of the monotonic trend used in the Mann–Kendall analysis.

Table 2. Summary of Mann–Kendall trend results for effective number of breeders (N_b).

Analysis window	Years	n	Kendall's τ	Sen's slope (N_b/yr)	Lag-1 autocorrelation (r_1)	MK p-value	Modified MK p-value
Long-term	2007–2024	18	0.27	+4.40	0.473	0.11	0.34
Short-term	2017–2024	8	0.00	-0.22	0.018	1.00	1.00

Long-term (2007–2024)

Over the full monitoring period (2007–2024), N_b shows a weak upward tendency, but the evidence is not strong enough to conclude a statistically significant increase (Table 2). The Mann–Kendall test produced a positive Kendall's τ, but the trend was not significant, and Sen's slope indicates a modest average increase of approximately 4 N_b per year. Autocorrelation in the data series was moderate, so the autocorrelation-adjusted test was also not significant.

Although the estimated Sen's slope suggests a modest positive rate of change, the lack of statistical significance indicates that the observed pattern is not strong enough to be distinguished from background variability. The presence of moderate autocorrelation further reduces the effective number of independent observations, limiting the power to detect gradual change. As a result, the long-term analysis should be interpreted as indicating no clear directional trend, rather than evidence of recovery or decline.

Short-term (2017–2024; ~two generations)

Over the most recent two generations (2017–2024), N_b appears stable with no detectable monotonic trend (Table 2). The Mann–Kendall results indicate no directional change, and Sen's slope is near zero during the two-generation time period. Autocorrelation over this period is minimal, and the autocorrelation-adjusted test leads to the same conclusion.

Because the short-term window reflects contemporary reproductive output, the absence of a detectable trend suggests that recent breeder contribution has remained stable. This result is particularly informative because the short-term window is less influenced by early high-N_b years and more reflective of contemporary reproductive conditions. The negligible autocorrelation observed over this period indicates that annual values are largely independent, strengthening confidence in the conclusion that no short-term directional change is evident.

Effective Number of Breeders Conclusion

Because N_b is estimated from age-0 fish collected in fall, it is best interpreted as an index of breeder contribution and early cohort production rather than an estimate of the number of adult grayling in the basin. The long-term record shows a mild upward tendency, but interannual variability is high and autocorrelation reduces statistical power. In other words, autocorrelation reduces the power to detect trends, so weak non-significant results should be interpreted as equivocal rather than confirmatory. Several years have very high N_b (e.g., 2012–2014 and 2019), which can influence visual impressions of change; however, the MK framework evaluates whether values are consistently higher later in the record, not whether there are episodic spikes. The longer record hints that N_b might be improving, but the statistics say we cannot confidently rule out chance variation. Over the recent two-generation period, N_b looks stable (no clear up or down signal).

Overall, the N_b time series provides no statistical evidence of sustained improvement or deterioration in breeder contribution over either the long-term or short-term analysis windows. The distinction between weak positive tendencies and statistically significant trends is critical: apparent increases visible in plots are driven in part by episodic high years rather than consistent improvement through time. These findings underscore the importance of interpreting N_b as a genetic and reproductive indicator that complements, but does not replace, demographic and habitat-based assessments of population status.

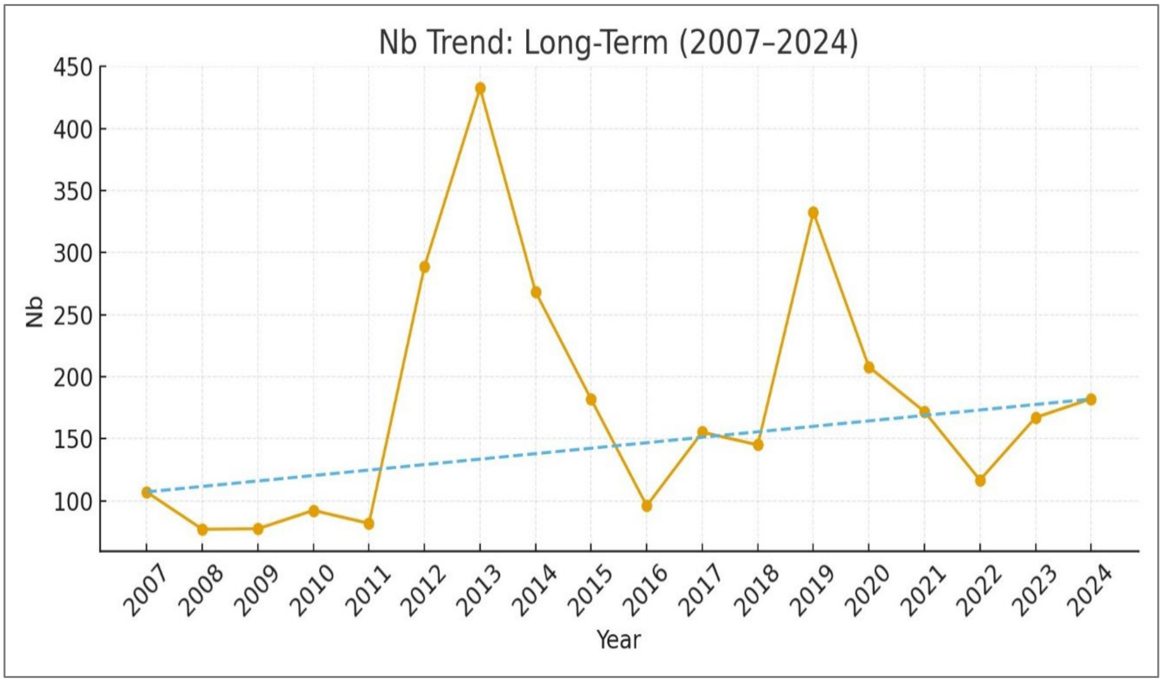

Figure 1. Long-term trend in effective number of breeders (Nb) for Big Hole River Arctic grayling, 2007–2024.

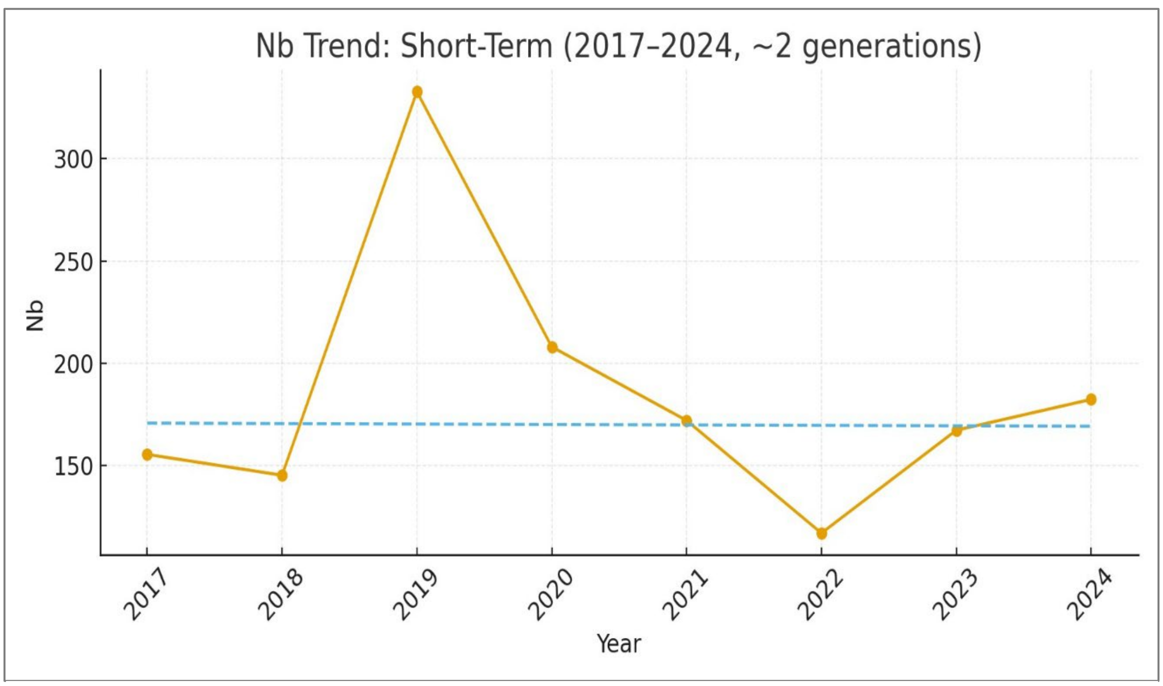

Figure 2. Short-term trend in effective number of breeders (Nb) for Big Hole River Arctic grayling, 2017–2024 (approximately two generations).

CPUE Trends for Arctic Grayling

This section describes how catch per unit effort (CPUE) data from Fall (September–October) single-pass electrofishing surveys were standardized, aggregated, and analyzed to evaluate directional trends in Arctic grayling abundance indicators in the Upper Big Hole River basin (2007–2024). This period was selected to evaluate whether CPUE shows sustained improvement during the CCAA era.

CPUE is the number of fish captured divided by sampling effort; here, effort is the length of stream surveyed. We report CPUE as fish per kilometer (fish/km), stratified by habitat (Mainstem Big Hole River vs Tributaries) and age class (young-of-year [YOY] vs Age-1+). Annual indices are computed with a design-consistent ratio-of-sums estimator (Σ fish ÷ Σ km) so that years with different reach lengths and site coverage remain comparable.

Data Source

We downloaded survey and inventory records from Montana Fish, Wildlife and Parks' (MFWP) Fisheries Information System (MFWP 2025). MFWP's database includes surveys conducted by MFWP staff and partner agencies and is updated on an ongoing basis. We analyzed Fall (September–October) single-pass electrofishing surveys in the Upper Big Hole River basin (mainstem and tributaries) for 2007–2024 to align with MFWP's CPUE reporting window and reduce seasonal bias. Reaches influenced by remote site incubators were excluded when present (e.g., Rock Creek and Governor Creek) to reflect "natural population" summaries.

MFWP's survey records include upstream and downstream endpoints but not reach length. We computed reach lengths by mapping the upstream and downstream endpoints of each survey on to the stream routes in the Montana Streams GIS layer (MFWP 2022) which serves as the hydrographic linear reference system for the Fisheries Information System. For each survey, we linked each survey record to a single stream feature in the Montana Streams GIS layer. For each survey, downstream and upstream coordinates (WGS84) were projected to NAD83 / Montana (meters), snapped to the appropriate stream route, and converted to along-route positions; the absolute difference between positions (converted to kilometers) was used as the survey length.

CPUE Calculation and Annual Aggregation

Following MFWP conventions (e.g., Magee et al. 2012), grayling were stratified into YOY and Age-1+. MFWP's YOY threshold (≤6.0 inches total length) was implemented as ≤150 mm (YOY) and >150 mm (Age-1+) to match millimeter reporting.

For each survey, we calculated CPUE as fish captured divided by reach length (km), yielding fish/km. To obtain a single annual CPUE value for each stratum × age class, we used a ratio-of-sums estimator: total fish captured across all fall surveys in that year and group divided by the total kilometers surveyed (Σ fish ÷ Σ km). This approach is equivalent to a length-weighted average of reach-level CPUE and avoids overweighting short reaches.

Trend Methods

We constructed annual fall CPUE indices separately for the Mainstem and Tributaries and for each age class (YOY and Age-1+). Years with missing fall surveys in a stratum were absent from that stratum's time series; we did not impute values. Because mainstem fall surveys were not conducted in 2019–2023, we report trends by stratum (Mainstem vs Tributaries rather than as a single basin-wide line) to avoid composition bias from changing survey coverage.

Monotonic trends were evaluated with Kendall's τ and the Theil–Sen slope (fish/km/year) with 95% confidence intervals. These non-parametric estimators are robust to outliers and distributional irregularities typical of electrofishing indices.

Results

Fall CPUE showed contrasting patterns between YOY and older fish (Table 3; Figures 3–4). In the Mainstem, YOY CPUE increased modestly and was statistically significant (Kendall's $\tau = 0.44$, $p = 0.042$; Theil–Sen slope = 0.123 fish/km/year, 95% CI 0.010 to 0.449). Mainstem Age-1+ showed no detectable trend ($\tau = 0.31$, $p = 0.164$; slope CI overlaps zero). In Tributaries, YOY CPUE showed no significant long-term trend ($\tau = 0.20$, $p = 0.308$), whereas Tributaries Age-1+ declined significantly ($\tau = -0.47$, $p = 0.008$; slope = −0.344 fish/km/year, 95% CI −0.668 to −0.111).

Table 3. Fall (September–October) CPUE trend results for Arctic grayling in the Upper Big Hole River basin, by stratum and age class (2007–2024).

Group	n years	Age Class	Median CPUE	Kendall's τ	p-value	Theil–Sen Slope	95% CI
Mainstem	13	YOY	0.37	0.44	0.042	0.123	(0.010, 0.449)
Mainstem	13	Age-1+	0.78	0.31	0.164	0.041	(-0.027, 0.146)
Tributaries	17	YOY	3.46	0.20	0.308	0.215	(-0.310, 0.674)
Tributaries	17	Age-1+	2.57	-0.47	0.008	-0.344	(-0.668, -0.111)

Trends are considered statistically significant at $\alpha = 0.05$ when Kendall's τ has $p < 0.05$ and the Theil–Sen 95% CI excludes zero.

The annual series reinforce these results. Tributaries show a conspicuous YOY pulse in 2019–2021 without a corresponding increase in Age-1+ CPUE (Figure 3), and Mainstem Age-1+ remains low with no clear improvement across sampled fall years (Figure 4). Overall, recruitment signals (YOY increases in some years/places) did not translate into higher fall abundance of older fish.

The implications for the CCAA era are straightforward. Across the monitoring record, fall CPUE provides no evidence of increasing abundance. Tributaries Age-1+ declined significantly, and Mainstem Age-1+ showed no positive trend. The significant Mainstem YOY increase is small in absolute terms and, without corroborating gains in Age-1+, suggests that many YOY do not survive (or are not recruited) into the Age-1+ population.

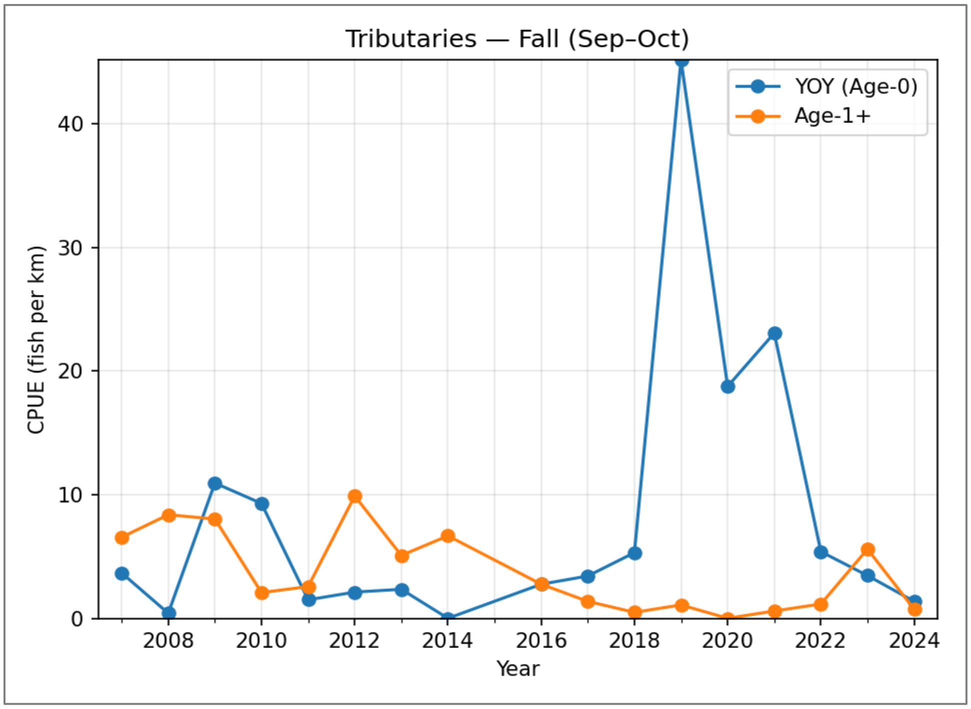

Figure 3. Tributaries—CPUE by age class Fall (September–October; 2007–2024).

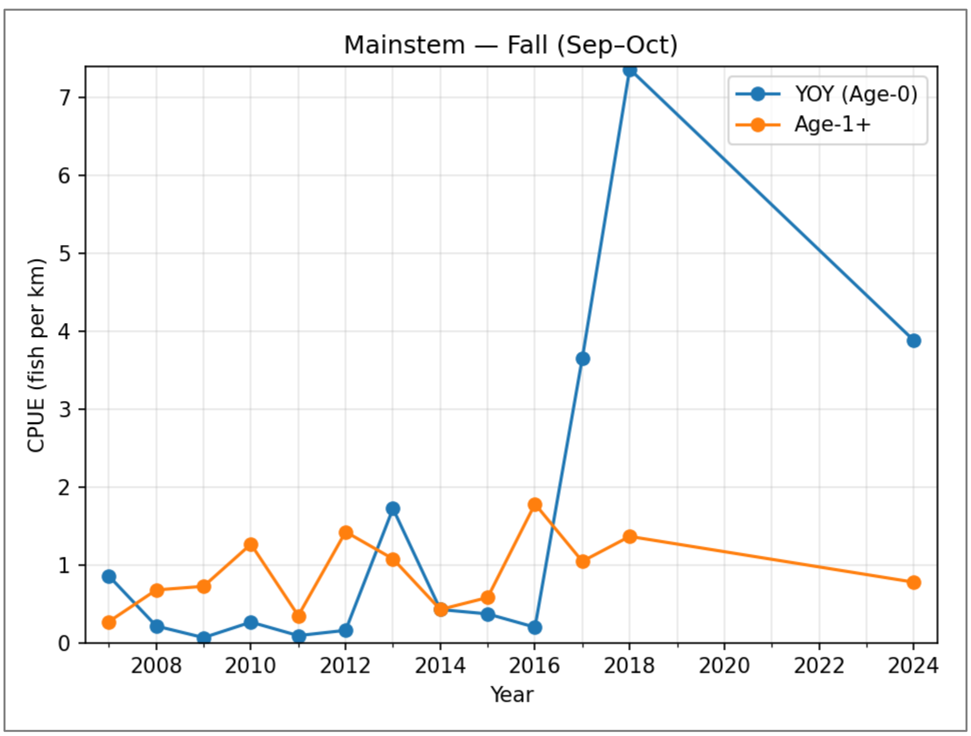

Figure 4. Mainstem Big Hole River—CPUE by age class Fall (September–October; 2007–2024).

CPUE Conclusions

Fall CPUE provides no evidence of improved abundance in the CCAA era. Mainstem Age-1+ shows no significant trend and remains low in absolute magnitude, while Tributaries Age-1+ declined significantly. Episodic YOY increases (including the 2019–2021 tributary pulse) have not carried forward into Age-1+ CPUE, suggesting that greater YOY production has not translated into sustained gains in older fish. Considered alongside N_b (which is inherently weighted toward YOY by design), the CPUE evidence similarly suggests that increased recruitment has not produced a durable increase in the abundance older age classes during the period evaluated. Overall, the combined pattern points to a bottleneck between YOY and older fish—consistent with limited survival and/or recruitment into the pool of future spawners. Because recovery ultimately depends on increasing the number of reproducing adults, the monitoring record indicates stagnation in older-fish abundance and implies that one or more limiting factors continue to constrain progression to maturity.

Reconciling N_b and CPUE

In the sections above, we evaluated two commonly used population viability indicators for Arctic grayling in the Big Hole River basin—CPUE from standardized field surveys and N_b estimated from genetic data—to understand why they imply different trajectories through time. Rather than treating these metrics as competing estimates of the same quantity, we interpret them as complementary indicators that are sensitive to different life stages and different sources of variation. This framing is particularly important when assessing potential responses to management actions such as implementation of the CCAA. Age-0 abundance can respond rapidly while recruitment to older age classes may change more slowly or remain constrained by persistent bottlenecks.

N_b and CPUE, in the context of Big Hole River Arctic grayling, differ fundamentally in what they measure. N_b and CPUE are derived from the same monitoring program, but not from identical sampling frames: N_b is estimated from age-0 fish collected during a fall sampling event, whereas CPUE integrates multiple ages and surveys. N_b reflects the effective number of breeders contributing genetically to the sampled cohort (i.e., breeder participation and variance in reproductive success) and early survival of recruits to the fall sampling window. CPUE, in contrast, is an index of the abundance of all age classes based on survey effort. As a result, N_b can change abruptly with increased spawning and/or survival of recruits to their first fall, while CPUE often changes more gradually and can remain flat when survival beyond the first year is limited. N_b is most sensitive to how many adults successfully produced offspring that year. It can increase quickly when reproduction and early survival improve—even if the number of fish that recruit to older age classes does not increase. Divergence between the metrics is therefore expected from life-stage differences and can be amplified by within-season sampling variability and subsampling choices.

Sampling alignment and implications.

Genetic samples used to estimate N_b are collected only during fall surveys, and in practice are taken during a single fall visit even when multiple fall surveys occur within a season. As a result, N_b reflects the genetic composition of age-0 fish available during a specific fall sampling event. CPUE, in contrast, is calculated from capture data across fish of multiple ages and may integrate multiple survey events in the spring and fall. This difference in sampling frame is important because within-season field conditions and fish availability can vary among fall visits. Accordingly, some divergence between N_b and CPUE may reflect not only life-stage sensitivity (age-0 recruitment versus multi-age standing stock), but also the greater event-specific sensitivity of N_b to the particular fall sampling occasion used for genetic collection.

Why Are They Different?

Early CCAA actions could plausibly improve access to and the quality of spawning habitat, early rearing habitat, or localized flow/temperature conditions in ways that increase the number of breeders that successfully contribute to a cohort. That would show up quickly in YOY-based N_b. But translating that into more juveniles/adults requires survival through summer low-flow/heat windows and predation/competition. If those bottlenecks didn't improve (or worsened due to climate/flow/temperature), CPUE could stay flat even while N_b increases. The divergence is consistent with an early-life-stage response that does not fully propagate to older, age classes. In that scenario, N_b captures an increase in successful breeding and early recruitment, while CPUE remains constrained by survival bottlenecks later in the life cycle.

Some divergence likely reflects the fact that the two indicators are vulnerable to different non-biological sources of variation. N_b is sensitive to cohort genetic structure and the composition of sampled juveniles, whereas CPUE is sensitive to capture probability and field conditions.

Taken together, the metrics are consistent with a system where recruitment can improve in some years—potentially aligning with early management implementation—yet overall abundance remains constrained by later-life-stage survival and habitat limitations that are not fully addressed by recruitment alone.

The observed divergence—N_b suggesting a marginal increase in abundance or productivity in the early years of CCAA implementation while CPUE does not show a comparable increase—is consistent with a scenario in which early-life-stage processes improved without a proportional increase in older fish. In this interpretation, N_b is capturing increased contributions of breeders and early recruitment, whereas CPUE remains constrained by post-recruitment bottlenecks such as overwinter survival, summer low-flow and high-temperature stress, habitat fragmentation, and other factors that determine whether early recruits survive and enter older age classes. Under these conditions, a recruitment signal can be real and meaningful without immediately producing a detectable increase in basin-wide abundance.

N_b derived from YOY reflects the effective number of breeders that contributed to that cohort, not the total number of adults in the population. Because N_b and CPUE emphasize different life stages and time scales, we next examine genetic diversity to evaluate longer-term resilience and whether USFWS's genetics-based inferences are consistent with the demographic record summarized above.

Genetic Diversity

This chapter summarizes how USFWS assessed genetic diversity for the Big Hole River population and provides a focused critique of USFWS's assumptions and conclusions regarding extinction risk (USFWS 2020). The Service and MFWP routinely report effective population size (N_e) alongside expected heterozygosity (H_e) and allelic richness (A_r). Here we treat these three metrics as complementary: N_e sets the time scale of genetic drift and inbreeding, H_e reflects average neutral genetic diversity, and A_r is sensitive to the loss of rare alleles that contribute to long-term adaptive potential. This section does not reanalyze raw genetic data; it evaluates whether the inferences drawn in USFWS (2020) follow from the analyses and data summaries cited. Below, after a brief primer on N_e, H_e, and A_r, we make four points: (1) trend tests have low power; (2) the genetic sampling chronology weakly isolates post-CCAA conditions; (3) heterozygosity-only projections understate allelic loss; and (4) N_e values near ~300 reduce concern about immediate inbreeding but do not, by themselves, demonstrate long-term adaptive capacity.

Genetic Diversity Metrics (N_e, H_e, A_r)

N_e is a measure of the population's genetic 'size' - the size of an idealized population that would experience the same amount of genetic drift as the population being studied. In practice, N_e is usually smaller than the number of fish on the landscape and can vary through time with changes in breeding success, age structure, and connectivity. Expected heterozygosity (H_e) measures the probability that two gene copies drawn at random from the population are different; H_e tends to change slowly and can remain stable even when rare alleles are being lost. Allelic richness (A_r) is a standardized count of the number of alleles per locus; A_r is more sensitive to bottlenecks and drift because rare alleles are lost first. For long-term resilience, this distinction matters: populations can show little change in H_e while still losing alleles that support future adaptive capacity (Allendorf et al. 2024).

USFWS' Analysis of Genetic Diversity

USFWS (2020) summarized multiple genetic datasets spanning several decades (e.g., microsatellite diversity metrics and estimates of N_b and N_e). For Big Hole River, USFWS emphasized two conclusions: (1) H_e and A_r were "relatively high" and had "remained stable" in recent sampling periods, and (2) genetic variation was expected to be lost slowly going forward (USFWS 2020, pp. 64–65). To support the second conclusion, USFWS referenced an N_e estimate on the order of ~300 and applied standard drift expectations to state that the Big Hole River population would lose less than ~10% of heterozygosity over the next 50 generations (~200 years) if that N_e is maintained (USFWS 2020, p. 64; citing Kovach et al. 2019).

USFWS (2020) acknowledged that one analysis documented a decline in allelic richness in the Big Hole River population, but characterized most of the decline as occurring in the 1990s to early 2000s, with "no trend in more recent years," and emphasized that heterozygosity remained stable (USFWS 2020, p. 64). USFWS then treated the overall picture as one of stable, relatively high genetic diversity and used this as part of a broader "line of evidence" that the population is more robust and resilient than in the past (USFWS 2020, p. 65).

N_e is a key metric in MFWP monitoring and USFWS findings because it provides a common scale for interpreting H_e and A_r through time and sets expectations for the pace of genetic drift and inbreeding. It is also the parameter USFWS uses to justify its '<10% heterozygosity loss over 50 generations' statement for the Big Hole River (USFWS 2020, p. 64; citing Kovach et al. 2019).

Interpreting this N_e-based projection works best if we separate short-term inbreeding risk from long-term adaptive capacity (Jamieson & Allendorf 2012). An N_e on the order of ~300 is above commonly cited near-term inbreeding guidelines (often framed around 50–100), supporting the narrow conclusion that inbreeding is unlikely to be an immediate driver of extinction risk (Franklin & Frankham 1998; Frankham et al. 2014). But it falls below classic long-term benchmarks (~500) and below more precautionary proposals (~1,000), so it does not by itself demonstrate maintenance of adaptive potential (capacity to respond to future environmental change; often reflected in allelic richness and retention of rare alleles) over the foreseeable future (Franklin & Frankham 1998; Jamieson & Allendorf 2012; Frankham et al. 2014).

A common shorthand for this tiered logic is the '50/500' guideline: N_e around 50 to reduce near-term inbreeding risk, and N_e around 500 to reduce longer-term loss of variation from genetic drift. These values were never meant to function as hard extinction thresholds; they provide genetic context that should be interpreted alongside demography and threats in a broader viability assessment (Jamieson & Allendorf 2012). More recent reviews suggest both benchmarks can be too low in wild populations and propose revised guidance closer to 100/1,000 (Frankham et al. 2014).

What Studies Say About Genetic Diversity

USFWS's 2020 finding relies on the research of Kovach et al. (2019). Kovach et al. (2019) quantify temporal trends in allelic richness (A_r) using simple linear regression, using the midpoint year of each multi-year collection as the predictor variable (e.g., 1987.5 for the 1987/1988 collection). Kovach et al. (2020) describe the same approach. Using all available time points, Kovach et al. (2019) report a weak but statistically significant decline in A_r over time for the Big Hole River ($P = 0.03$), but they note that the trend becomes non-significant when the first observation is removed ($P = 0.11$) and interpret this as A_r being "relatively stable" through the 2000s and 2010s. Kovach et al. (2020) present the same general pattern: a weak overall decline ($P = 0.053$) that becomes non-significant after removing the first observation ($P = 0.294$), again describing A_r as "relatively stable" in the 2000s and 2010s.

The key limitation, however, is that this interpretation rests on a sensitivity check with very few data points. Kovach et al. (2019 and 2020) used linear regression to quantify temporal trends between five A_r observations taken from the mid-1980s through the mid-2010s. Once the first observation is dropped, only four points remain (degrees of freedom = 2), so the p-value becomes highly sensitive to which subset is analyzed and provides limited leverage for a strong "trend vs. no trend" conclusion.

More importantly, a "non-significant" result should not be treated as proof of "no change" or "stability." A non-significant p-value means that—given the model, assumptions, and the data available—the analysis does not provide strong statistical evidence for a trend in that reduced dataset; it does not demonstrate that the true trend is zero. Statistical guidance cautions against turning a single p-value threshold into a yes/no scientific conclusion, especially when the dataset is small and tests are low power (Wasserstein and Lazar 2016). This general concern is amplified here because the reduced Big Hole time series is very short (four points after removing the first observation), and standard references note that small samples often yield large uncertainty and low power, making "no significant trend" a fragile basis for strong inference (National Research Council 2011).

Limited Alignment with the CCAA Framework

Another problem is that the sample years do not line up cleanly with a CCAA effectiveness claim. The Big Hole River CCAA was issued in 2006, yet the "drop the first point" subset that drives the "stable" interpretation still includes midpoints in the mid-1990s and mid-2000s—years that are wholly or largely pre-CCAA. That regression therefore cannot isolate a post-2006 pattern, so it is a weak basis for attributing "stability" to the CCAA.

Why Heterozygosity Alone Understates Allelic Loss

USFWS also argues that loss of genetic diversity is not an immediate threat, citing projections that expected heterozygosity (H_e) would decline by <10% over the next ~50 generations if N_e remains near current levels. That projection is informative, but it does not capture the most sensitive component of genetic erosion. Conservation genetics theory and empirical studies show that allelic richness declines more rapidly than heterozygosity during bottlenecks and under genetic drift because rare alleles are lost first; as a result, populations can show little change in heterozygosity while still losing alleles that contribute to long-term adaptive potential. This point is highlighted in modern synthesis treatments (Allendorf et al. 2024) and foundational bottleneck literature (e.g., Nei et al. 1975; Luikart et al. 1998). Consistent with this expectation, Kovach et al. (2019, 2020) report declines in allelic richness even when heterozygosity remained stable. Taken together, H_e-based projections alone likely understate ongoing or future erosion of allelic diversity, and the <10% figure should be treated as conditional reasoning rather than a guarantee. The point is not that the population is on the verge of genetic collapse; it's that

‘<10% H_e loss’ is an incomplete framing for long-term resilience because it discounts faster erosion of allelic diversity.

Genetic metrics do not replace demographic evidence, but they inform our interpretation of long-term resilience and adaptive capacity in a changing environment; the synthesis that follows integrates demographic patterns with the genetic context summarized above to assess overall viability.

Viability Synthesis

N_e provides context for expected rates of genetic drift, but it does not substitute for demographic evidence of a stable or recovering population. N_e values near ~300 suggest the population is not on the verge of rapid inbreeding-driven collapse, yet they do not demonstrate long-term adaptive capacity or offset the lack of sustained improvement in N_b and CPUE under persistent instream flow and temperature stress.

Taken together, N_b and CPUE point to a persistent constraint in survival beyond the first year. N_b is a recruitment-weighted genetic indicator because it is estimated from a single cohort of young fish, whereas CPUE more directly reflects the abundance of older age classes across habitats. Over the monitoring period, N_b shows at most weak or episodic improvement, but CPUE shows no corresponding basin-wide increase and declines in tributaries. This divergence is consistent with years of improved early recruitment that do not translate into increased retention to older age classes, which would occur if post-recruitment survival remains constrained by summer/fall low flows and high temperatures, habitat fragmentation, or other chronic stressors.

A credible recovery signal would be sustained and concordant across multiple independent indicators—particularly those tied to survival toward maturity. Instead, the monitoring record contains mixed or flat signals that do not support a conclusion of basin-wide improvement or reduced extinction-risk based on abundance indicators alone. This does not demonstrate that individual management actions were ineffective; it indicates that the available evidence does not justify confident “recovery” inference. The same standard of evidence should be applied when interpreting genetic metrics, especially when forward-looking statements rely primarily on heterozygosity rather than metrics more sensitive to allelic loss.

Key Threats: Instream Flow and Temperature

The demographic and genetic evidence above points to a system where early recruitment occurs, but survival into older age classes is consistently constrained. The most plausible and repeatedly cited basin-scale mechanisms are summer/fall low flows and high temperatures, which compress habitat and increase physiological stress during the period of highest vulnerability. Accordingly, this chapter evaluates whether CCAA flow targets and temperature benchmarks align with grayling biological needs and whether trends indicate meaningful improvement. This focus mirrors the drought-response trigger framework adopted in the 1990s, which linked management response to warm-season flow deficits and temperatures around the 70°F stress threshold (Big Hole River Watershed Committee 1997).

Seasonal Instream Flow Targets

This chapter evaluates whether streamflow conditions relevant to Arctic grayling habitat are improving, stable, or declining in the Big Hole River under the CCAA framework. We use the CCAA's seasonal minimum instream-flow targets for each management segment (spring: April–June; summer/fall: July–October) as a consistent benchmark for habitat adequacy. Using daily discharge records from the gage network, we summarize each year as the percent of days meeting the seasonal target, which translates complex hydrographs into an intuitive measure of compliance with minimum objectives. We then test for long-term trends in seasonal compliance for each segment using nonparametric methods (Mann–Kendall with Theil–Sen slope) appropriate for hydrologic time series. Finally, for Segments C and D—where the gage record supports it—we add a before/after comparison to evaluate whether compliance shows a step change across the CCAA era.

For analytical purposes, the CCAA specifies seasonal flow targets for each of the five management segments and a program goal of meeting targets on at least 75% of days in each season (Table 4 and Figure 5). Spring targets correspond to the upper inflection points from wetted-perimeter analyses (intended to represent more biologically optimal habitat conditions), whereas summer/fall targets correspond to the lower inflection points (intended to maintain minimum connectivity during the irrigation season). For example, at Wisdom (Segment C), the spring target is 160 cfs and the summer/fall target is 60 cfs; farther downstream at Mudd Creek Bridge (Segment D), the targets are 350 cfs and 100 cfs, respectively.

Table 4. Seasonal minimum flow targets for the management segments and monitoring locations adopted by the CCAA (USFWS 2006).

Management Segment	Location of Discharge Monitoring	Minimum Flow Targets (cfs)	
		April–June	July - Oct
A	Miner Creek Road	60	20

B	Little Lake Creek Road	100	40
C	Wisdom	160	60
D	Mudd Creek Bridge	350	100
E	Dickie Bridge	450	170

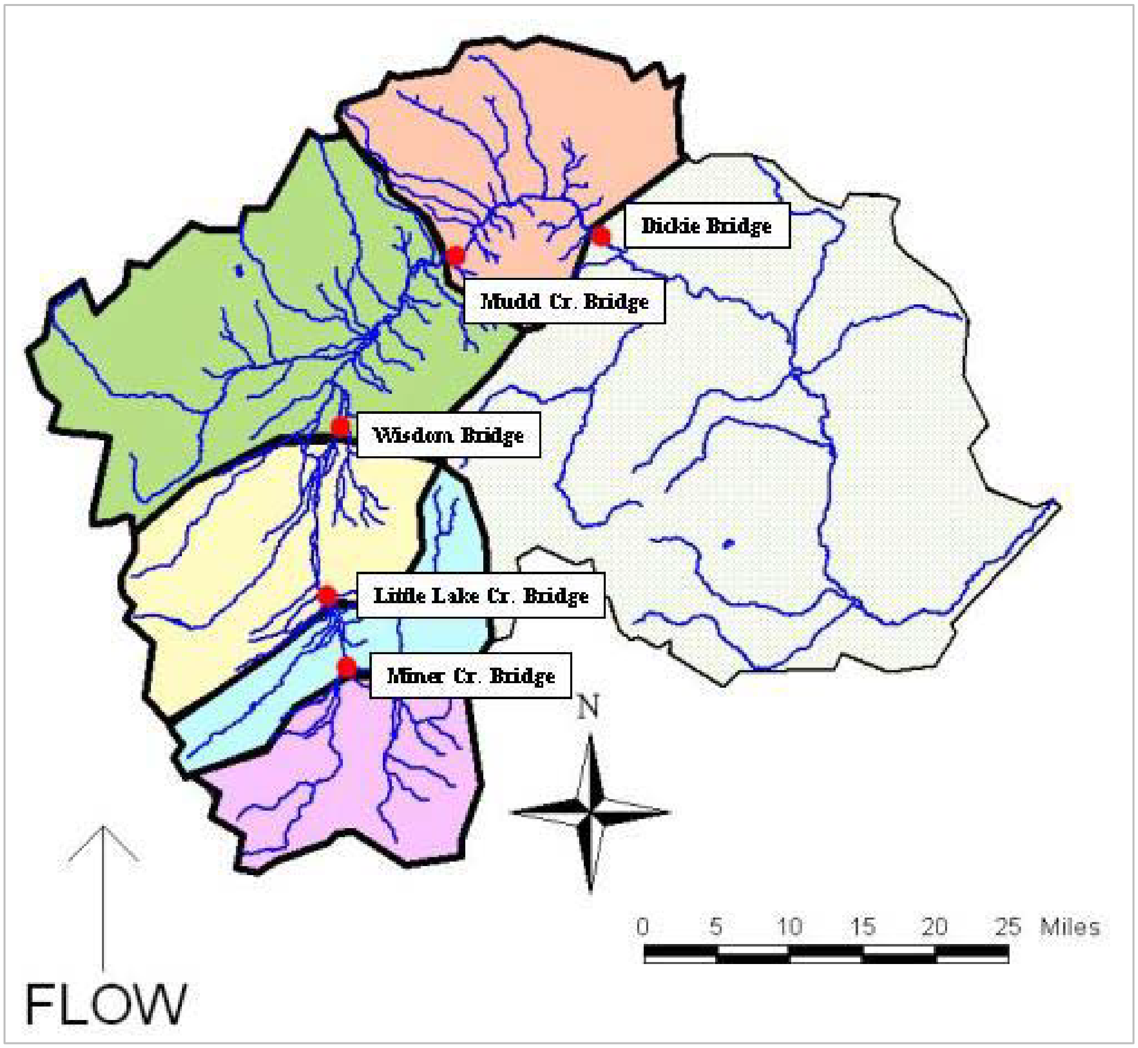

Figure 5. Big Hole River Management Segments (CCAA 2006).

Scientific Basis and Limitations of Targets

The Big Hole River CCAA bases its instream flow targets on the same Montana "wetted perimeter" methodology that Montana Fish, Wildlife & Parks (MFWP) developed in the early 1980s and that Leathe and Nelson formally summarized in 1986[1]. The Agreement uses the same framework and logic as the original Montana instream flow evaluations. However, the way the CCAA converts that science into specific minimum flow targets—especially for summer and fall—reflects a clear compromise between aquatic habitat requirements and irrigation demands. Leathe and Nelson (1986) explicitly state that flows at the lower inflection support only a low level of habitat potential and may not be adequate for rare or sensitive species, and they recommend using flows at or near the upper inflection point for high-value streams and waters with species of special concern, including Arctic grayling. The CCAA instead reserves upper-inflection protection for spring and intentionally accepts lower-inflection conditions in summer and fall, when baseflows are lowest and fish are most vulnerable.

A U.S. Fish and Wildlife Service habitat-suitability assessment for Arctic grayling (Hubert et al. 1985) independently reinforces this concern about summer/fall flows in the upper Big Hole River basin. Drawing on earlier studies, the report notes that reduced streamflows have already been identified as a factor limiting grayling production in the drainage, underscoring that low-flow conditions are not just a theoretical risk but a demonstrated constraint on the population. Taken together, these lines of evidence support close scrutiny of summer/fall compliance trends. Persistent summer/fall deficits—especially when coupled with warm temperatures—are consistent with declining habitat availability during the period when grayling are most vulnerable, and minimum flows based on lower-inflection criteria may provide less protection than envisioned in the original instream-flow guidance for sensitive species.

Stream Gages and Flow Data

The hydrologic analysis of Big Hole River instream flows integrates data from a network of stream gages that span the five management segments (A–E) identified in the CCAA (Table 4, Figure 5). Each segment corresponds to a distinct reach of the river with unique hydrologic and ecological characteristics. Collectively, these gages provide the long-term discharge records needed to assess water availability, evaluate instream-flow compliance, and understand seasonal and year-to-year variability across the upper basin.

The stream gages for segments A, B, and E are managed by the Montana Department of Natural Resources and Conservation (DNRC). Although DNRC currently operates these sites, they rely on legacy U.S. Geological Survey (USGS) data for part of their historical

---

[1] The Big Hole River CCAA incorrectly cites the document as Leathe and Nelson 1989. The publication date of the document is December, 1986.

record. During the first few years that DNRC operated these gages, the USGS continued to host the data on its website. However, due to data-quality concerns, the USGS determined that the records did not meet their publication standards and chose not to post the data for public use. As a result, the records for these segments contain 2–4-year gaps in the early years of operation.

The Segment A gage (DNRC site 41D 02000; legacy USGS site 06023800) captures upper-basin hydrology near the river's headwaters. The Segment B gage (DNRC site 41D 05000; legacy USGS site 06024020) records early snowmelt responses and irrigation influences in the high-elevation Jackson area. The Segment E gage (DNRC site 41D 08000; legacy USGS site 06024580) represents the transition toward the lower basin and integrates flows from several major tributaries before the river enters the canyon reach.

In contrast, Segments C and D are managed directly by the USGS and provide the most continuous and complete hydrologic records in the system. The Segment C gage (USGS 06024450) has operated for several decades and delivers high-quality mid-basin discharge data useful for long-term trend and temperature analyses. The Segment D gage (USGS 06024540) captures flows in a reach strongly influenced by groundwater interaction and upstream irrigation diversions. Both USGS gages benefit from consistent methods, rigorous calibration, and high data reliability, making them essential reference points for basin-wide comparisons.

Together, the DNRC and USGS gage records provide a reasonably long and spatially extensive dataset for evaluating instream-flow targets, drought response, and hydrologic trends relevant to Arctic grayling conservation. The overlap between DNRC-managed and legacy USGS records creates multi-decadal time series that support detection of long-term changes attributable to climate variability, land use, and management actions implemented under the CCAA.

Long-term Trends for Segments A–E

We focused our long-term trend analysis on the years following the implementation of the CCAA. Evaluating long-term trends in seasonal compliance with CCAA instream flow targets provides a system-wide assessment of whether hydrologic conditions have improved, declined, or remained stable during the years since the agreement was implemented. Because each management segment has a different monitoring record and different operational history, trend analysis offers a consistent way to understand how streamflow performance has evolved over time, independent of changes in gage ownership, data gaps, or the timing of local conservation actions. This analysis focuses on the annual percentage of days meeting the CCAA flow targets for Spring (April–June) and summer/fall (July–October), the two biologically critical periods for fluvial Arctic grayling.

To quantify trends, we used the Mann–Kendall (MK) test and the Theil–Sen slope estimator, a pair of nonparametric tools specifically suited to hydrologic time series. The MK test evaluates whether the compliance time series shows a statistically significant upward or downward trend, while the Theil–Sen slope estimates the magnitude of change in percentage-points per year. Together, these approaches provide a robust, transparent evaluation of how instream flow performance is changing across all management segments with sufficient post-CCAA data. We present the results of the MK test and Theil–Sen slopes in Table 5 and Figures 6 & 7 below.

Table 5. Summary of Mann–Kendall Trend Test Results and Theil–Sen Slope Estimates. Negative slopes indicate decreasing seasonal compliance with instream flow targets.

Segment	Season	Years	p-value	Theil–Sen slope (% points per year)
A	Spring (Apr–Jun)	2010–2025	0.086	-0.943
	Summer/fall (Jul–Oct)	2009–2025	0.161	-2.563
B	Spring (Apr–Jun)	2010–2025	0.363	-0.581
	Summer/fall (Jul–Oct)	2009–2025	0.041	-2.937
C	Spring (Apr–Jun)	2006–2025	0.330	-0.680
	Summer/fall (Jul–Oct)	2006–2025	0.080	-2.168
D	Spring (Apr–Jun)	2006–2025	0.056	-0.375
	Summer/fall (Jul–Oct)	2006–2025	0.173	-1.129
E	Spring (Apr–Jun)	2010–2025	0.598	0
	Summer/fall (Jul–Oct)	2009–2025	0.008	-3.252

As stated above in the Methods section, The MK test assumes the yearly values are independent. If there’s strong serial autocorrelation, the MK p-values can look more significant than they really are. Confirmation of low serial dependence increases confidence that the MK p-values reflect true long-term patterns rather than temporal artifacts. We evaluated lag-1 autocorrelation in the seasonal compliance series and found low to moderate autocorrelation. These values were not large enough to require adjusting the variance or p-values of the Mann–Kendall trend test.

Figure 6. Annual percentage of days meeting spring (April–June) instream-flow targets for management segments A–E of the upper Big Hole River, with nonparametric trend lines based on Theil–Sen slope estimates and Mann–Kendall tests.

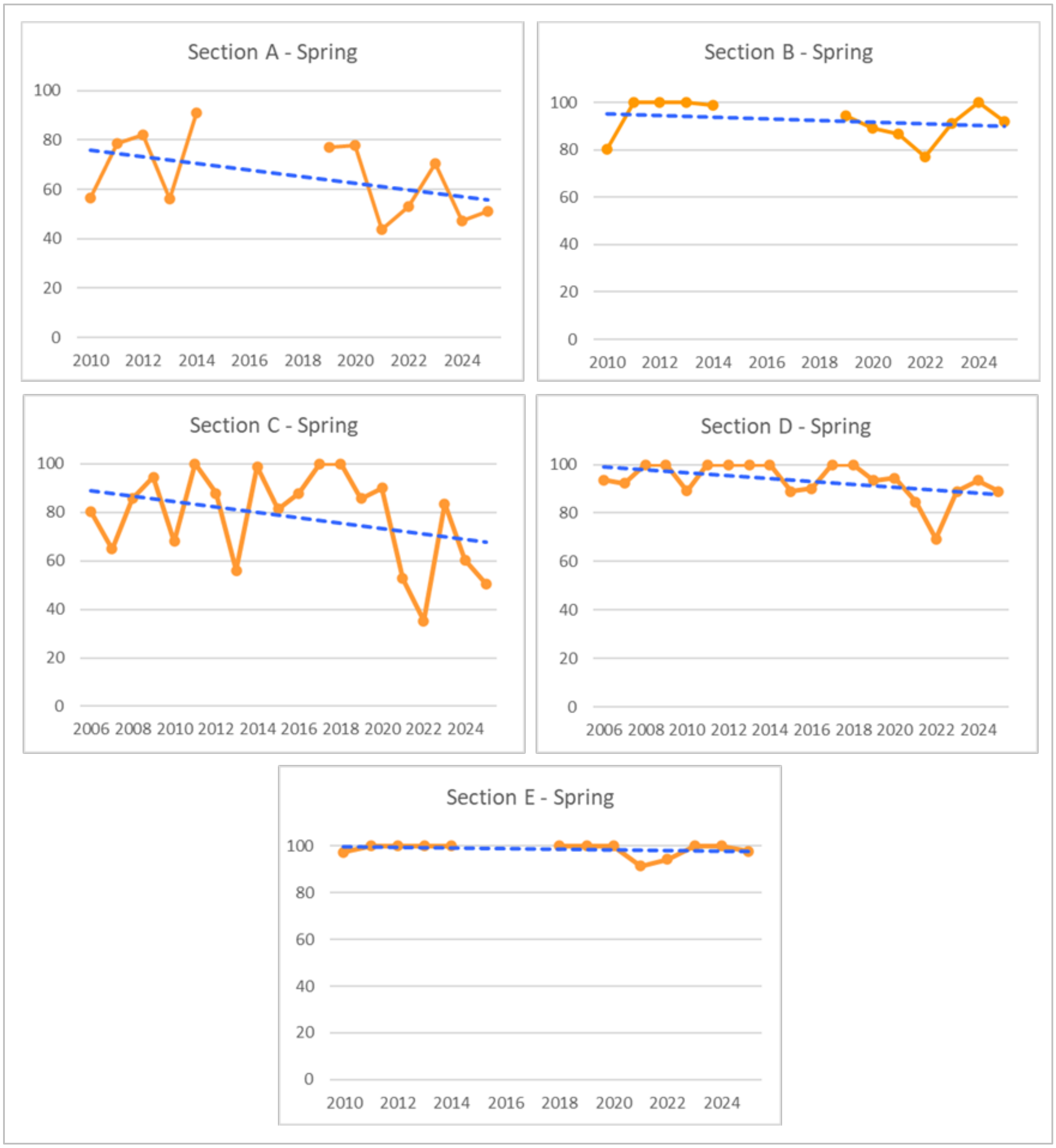

Figure 7. Annual percentage of days meeting summer/fall (July–October) instream-flow targets for management segments A–E, illustrating consistently lower compliance and stronger negative trends than in spring.

Segment A

In Segment A, spring instream flow conditions appear to be slowly declining and summer/fall conditions have largely failed to meet instream flow goals. Spring flows often meet the 60 cfs target, but performance has slipped over time. The trend is not statistically significant, but the negative Theil–Sen slope (–0.94 percentage points per year) shows a gradual decline in the likelihood of meeting the target. Over the period of record, spring compliance averages about 65%. Summer/fall performance is weaker with compliance averaging about 56%, and the trend showing a stronger decline (–2.56 percentage points per year). In the most recent seven years, compliance is consistently below 75%. Although the Mann–Kendall p-value is not significant, the last seven years fall well below the CCAA goal of meeting targets more than 75% of days[2] after the first 10 years and more frequently in the long-term. Overall, the pattern—especially in summer/fall—indicates declining summer/fall reliability in meeting flow targets, which can increase the likelihood of low-flow conditions during the most sensitive period for grayling.

Segment B

In Segment B, spring instream flow conditions are stable but summer/fall conditions have not met flow targets and appear to be deteriorating. Spring flows (100 cfs target) remain strong. Mean compliance (~92%) is high and shows no statistically detectable trend. Spring targets are consistently achieved in Segment B, indicating stable spring flow performance over the period of record. Summer/fall compliance is much lower and shows clear deterioration. Mean compliance (~70%) is below the >75% CCAA goal, and the MK test shows a statistically significant downward trend ($p = 0.041$) with a decline of nearly 3 percentage points per year. Recent years show very low compliance. The significant downward trend indicates summer/fall targets are being met less often, consistent with increasing risk of low-flow conditions during summer/fall.

Segment C

In Segment C, neither the spring nor the summer/fall instream flows are meeting the goals of the CCAA. Spring instream flows (160 cfs target) show moderate but declining performance. Compliance averages about 72%, and although the Mann–Kendall trend is not statistically significant, the Theil–Sen slope (about -0.7 percentage points per year) indicates a gradual weakening of spring conditions over the 2006–2025 period. Spring flows often approach the CCAA threshold of meeting targets at least 75% of days, but performance varies markedly across years, and dry springs depress compliance. Taken together, spring performance in Segment C is frequently near—but not reliably above—the 75% benchmark, underscoring sensitivity to dry spring conditions.

Summer/fall performance is substantially weaker and compliance with flow targets averages about 43%. The Theil–Sen slope (-2.17 percentage points per year) and a marginal MK p-value (0.08) suggest a likely decline in summer/fall reliability of meeting the target, even though the trend does not meet the conventional 0.05 threshold.

[2] The CCAA includes the following goal: “Flow targets met 75% of time by year 10 & more frequently thereafter as the more complex site-specific plans reach full implementation.”

Segment D

In Segment D, spring instream flows are generally good but summer/fall flows are well below the threshold. Spring flows (350 cfs target) remain relatively strong but show signs of weakening. Mean compliance is about 78%, which is just above the 10-year CCAA objective. The Theil–Sen slope (about –0.4 percentage points per year) and borderline MK significance (p = 0.056) indicate a gradual downward trend in spring flow performance. The emerging downward signal suggests increasing sensitivity to dry springs and a reduced margin above the 75% benchmark.

Summer/fall flows are considerably more limited and compliance with the threshold averages about 52%. The Theil–Sen slope (–1.13 percentage points per year) indicates a steady decline, though the MK trend is not statistically significant. Recent low-compliance years reinforce a pattern of worsening summer/fall low-flow conditions in Segment D.

Segment E

In Segment E, we see a similar pattern where the summer/fall instream flows are the limiting factor. Spring flows (450 cfs target) consistently meet or exceed the CCAA target. Mean compliance is about 98%, and no trend is evident. Spring conditions in Segment E appear stable and adequate for grayling habitat. Summer/fall performance is declining sharply. Although earlier years achieved high compliance, recent years fall far short of the 170 cfs target. Mean compliance (~73%) is below the 10-year CCAA objective, and the trend shows a statistically significant and substantial decline (–3.25 percentage points per year; p = 0.008). Some recent years fall below 40% compliance with the minimum instream flow threshold. These results show a marked reduction in the frequency of meeting the 170 cfs summer/fall target in recent years, consistent with increasingly severe summer/fall low-flow conditions in the lower Big Hole River.

Summary of Long-term Trends

Across all five management segments, spring (Apr–Jun) flows generally remain closer to CCAA instream-flow objectives than summer/fall (Jul–Oct) flows, but there are emerging signs of weakening spring performance in several segments. Spring trend tests show slight to moderate negative slopes in Segments A, B, C, and D, with the strongest evidence of decline in Segment A and a borderline decline in Segment D. Segment E shows no detectable spring trend. Although most spring trends do not meet the standard significance threshold ($p \leq 0.05$), the consistency of negative slopes across four segments suggests the pattern is not purely random year-to-year variation. Overall, spring conditions are often adequate, but they do not appear uniformly stable and may be gradually eroding in parts of the system.

In contrast, summer/fall compliance shows consistent declines across the upper basin. All five segments exhibit negative trends for summer/fall, with especially steep declines in Segments A, B, C, and E. Statistically significant or near-significant declines occur at multiple gages (notably Segment B and Segment E), and even where p-values exceed 0.05, the direction and magnitude of the slopes point toward increasing summer/fall flow deficits over time. Because these declines occur during the hottest and driest months—when habitat availability and connectivity are most constrained—they are the most consequential for Arctic grayling.

Before–After Trend Analyses: Segments C & D

Segments C and D are the only management segments with records long and continuous enough to support a direct comparison of instream-flow performance before and after implementation of the CCAA. Segment C uses the USGS 06024450 record (1988–2025) and Segment D uses the USGS 06024540 record (1998–2025). These before/after analyses complement the trend assessments by addressing a different question: Did adoption of the CCAA lead to measurable improvements in meeting instream-flow targets? For Segment C, we define the pre-CCAA period as 1988–2005 and the post-CCAA period as 2006–2025; for Segment D the pre-CCAA period is 1998–2005 and the post-CCAA period is 2006–2025. Because the pre-CCAA period is shorter for Segment D (1998–2005), before/after tests for that segment have less power to detect modest changes. We evaluate targets for spring (April–June) and summer/fall (July–October).

We used two complementary nonparametric approaches. The Mann–Whitney U test compares annual seasonal compliance (% days meeting the seasonal instream-flow target) between the pre- and post-CCAA periods to test whether the distribution shifted upward or downward. This test makes no distributional assumptions and is well-suited for hydrologic metrics with high interannual variability. We also fit logistic regression models using daily target attainment (meets target = 1, below target = 0) to test whether the odds of meeting the target differed between periods. We report results as odds ratios (post/pre) with robust standard errors clustered by year. Table 6 summarizes the before–after comparison of seasonal instream-flow compliance and the corresponding statistical test results for Segments C and D.

Table 6. Summary of before–after instream-flow performance and statistical test results for management Segments C and D.

Segment	Season	% Days Meeting Target (Before)	% Days Meeting Target (After)	MW p-value	Odds Ratio (Post/Pre)	Odds Ratio p-value
C	Spring	71%	80%	0.552	1.56	0.292
	Summer/Fall	38%	43%	0.412	1.23	0.592
D	Spring	90%	91%	0.682	1.22	0.682
	Summer/Fall	71%	71%	0.778	1.00	0.998

Note: Before/after periods: Segment C = 1988–2005 vs 2006–2025; Segment D = 1998–2005 vs 2006–2025.

Segment C

In Segment C, Post-CCAA years exhibited a higher spring mean percentage of days meeting the flow target (80% vs. 71%). Logistic regression also indicates that summer/fall days in Segment C are closer to the 60 cfs target after CCAA implementation (percent of days meeting the target increased from 38% to 43%). However, this level of performance remains well below the 75% threshold established for instream flow compliance.

The Mann-Whitney tests revealed no statistically significant pre/post differences ($p > 0.05$) in the spring and summer/fall percent of days meeting instream flow targets. However, both indicators were higher post-CCAA, particularly in spring. The non-significant p-values reflect high interannual variability and limited statistical power.

Logistic regression indicates that during spring, the odds of meeting the 160 cfs target post-CCAA were approximately 1.56 times higher than before. This improvement was not statistically significant ($p = 0.29$). For summer/fall, odds increased by a smaller margin (OR = 1.23, $p = 0.59$). Confidence intervals for both seasons include 1.0, implying that observed changes may be due to natural hydrologic variability.

Segment D

In Segment D, spring performance remains relatively high in both periods, with 91% of spring days meeting the 350 cfs target before the CCAA and 90% meeting the target afterward. This change is negligible, and the Mann–Whitney test confirms no statistically significant shift between pre- and post-CCAA spring conditions ($p = 0.682$). Logistic regression likewise indicates no meaningful change in the odds of meeting the spring target following CCAA implementation (Odds Ratio = 1.22, $p = 0.682$). Together, these results suggest that spring flows in Segment D were already strong before the CCAA and have remained generally stable, with no statistical evidence of improvement or decline attributable to the program.

Summer/fall conditions in Segment D show an even clearer pattern of stability across the pre- and post-CCAA periods. The percentage of days meeting the 100 cfs target was identical before and after CCAA implementation (71% in both periods), and the Mann–Whitney test indicates no difference between seasonal median compliance values ($p = 0.778$). Logistic regression also shows no change in the likelihood of meeting the summer/fall target (OR = 1.00, 95% CI: 0.32–3.09, $p = 0.998$). These results indicate that the CCAA did not produce a detectable shift in summer/fall flow reliability in Segment D. As with spring, summer/fall performance appears largely governed by broader hydrologic variability rather than by a discrete change associated with the transition to CCAA management.

Overall, the before–after analysis for Segment D shows no statistically detectable change in seasonal compliance following implementation of the CCAA. Spring conditions were high before the program and remain so afterward, while summer/fall performance has remained moderate and unchanged. Although these results do not indicate improvement, they also do not show deterioration attributable to the CCAA; rather, streamflow conditions in Segment D appear to reflect longer-term hydrologic patterns rather than an immediate management effect.

Summary of Before–After Trends

Logistic regression results for Segments C and D suggest that spring days may be somewhat more likely to meet the instream flow targets after CCAA implementation, but the estimated improvement is small and statistically uncertain. While the statistical evidence for a distinct spring "step change" is weak, spring flows in Segments C and D are typically closer to the desired performance standard than summer/fall flows.

For summer/fall in both Segments C and D, logistic regression results suggest at most small relative changes in the odds of meeting instream flow targets after the CCAA. However, the overall percent of days meeting the target remains below the 75% threshold, and short-term trend analyses show declining summer/fall conditions in recent years.

Instream Flow Conclusions

Overall, our analysis indicates that the Big Hole River CCAA hydrologic objectives for Arctic grayling are not being achieved consistently, especially in summer/fall. In the long-term analysis of spring flows several management segments fall short of the goal of meeting targets on at least 75% of days, and trend analyses show weak but consistently negative slopes in four of the five segments. These results suggest that spring conditions are not uniformly stable and may be gradually eroding in parts of the system. In the long-term analysis of summer/fall conditions the percentage of days meeting summer/fall instream-flow targets is well below the CCAA objectives in all five management segments and long-term trend tests show consistent negative slopes, with statistically significant or near-significant declines in multiple segments.

The before–after analyses for Segments C and D show no statistically detectable improvement in instream-flow performance associated with the implementation of the CCAA. Spring conditions in these segments have remained generally the same, with only small, statistically uncertain increases in the odds of meeting the target. Summer/fall performance in both segments remains well below the 75% objective, and logistic regression indicates no meaningful change in the likelihood of meeting the target between the pre- and post-CCAA periods. These findings indicate that, at the scale of the mainstem hydrograph, voluntary measures taken under the CCAA to date have not produced a clear, measurable improvement in seasonal compliance with instream-flow targets.

Finally, it is important to view these compliance patterns in light of how the CCAA instream-flow targets were set. The Agreement is formally grounded in Montana's wetted-perimeter methodology, which identifies lower and upper inflection points in riffle habitat as an envelope of acceptable low flows. Earlier technical guidance, including Leathe and Nelson (1986), explicitly recommended using flows at or near the upper inflection point for high-value streams and sensitive species such as fluvial Arctic grayling. The CCAA applies this upper-inflection guidance only in spring; summer/fall targets are based on the lower inflection point, a level that the original work acknowledged would provide only a low level of habitat and could be inadequate for rare or sensitive species. In practice, the river is failing to meet even these lower-bound summer/fall targets with increasing frequency.

Thus, the current management framework combines less-protective summer/fall targets with declining summer/fall compliance, a combination that is unlikely to sustain Arctic grayling habitat under current and projected climate and water-use conditions.

Taken together, these results indicate that the CCAA, in its current form, has not been sufficient to maintain or improve summer/fall flow conditions in line with its own instream-flow objectives, and that summer/fall flows both fall short of and are less protective than the minimums originally recommended for Arctic grayling streams. Addressing this gap may require a combination of stronger flow-enhancement measures, more protective summer/fall targets, and drought-mitigation strategies that explicitly account for increasing climate stress and irrigation demand in the upper Big Hole River.

Stream Temperature

This chapter evaluates instream water temperatures in the upper Big Hole River and the temperature criteria used in risk determinations. This chapter primarily synthesizes reported patterns and metrics; limited supplemental calculations (e.g., warming rates at USGS gages) are noted where relevant. This chapter is presented in two parts. Because flow and temperature operate together, the assessment first defines temperature criteria and indicators that reflect chronic ecological performance. The assessment then summarizes temperature patterns reported in the monitoring record and evaluates the implications for the CCAA.

Flow and temperature are linked stressors in the upper Big Hole. When summer/fall flows drop, usable habitat contracts, shallow margins warm faster, and thermal refuges can become disconnected or inaccessible. As a result, even modest long-term declines in summer/fall flow compliance can translate into disproportionate biological risk by increasing both the frequency and duration of warm-water exposure during the period when grayling are most vulnerable. For that reason, the next section evaluates temperature risk using criteria that reflect ecological performance rather than short-term laboratory lethality.

Thermal criteria in the scientific literature are typically expressed in °C, whereas the Big Hole monitoring reports often summarize exceedances in °F. In this report, we summarize monitoring results in the units used by the source reports and review biological thresholds in °C (with °F equivalents shown at key reference points).

Temperature Trends and Exceedances

This section synthesizes temperature patterns and exceedances as reported in agency monitoring documents; it does not reprocess raw temperature time series or re-express the monitoring record relative to the alternative criteria proposed in the previous section of this report.

Drawing on three decades of monitoring summarized in Early Recovery/Restoration reports, MFWP Arctic Grayling Monitoring Reports, and CCAA annual reports—along with basin-scale analyses by Sladek (2013) and Vatland (2015)—the available evidence does not indicate a measurable improvement in summer water temperatures in the upper Big Hole River during the CCAA era. Across all periods, the same spatial pattern recurs: relatively cool headwater and spring-influenced reaches contrasted with a persistent set of mainstem and tributary hot spots that frequently exceed 70°F (21.1°C) and, in some years, reach or exceed 77°F (25°C). Interannual variability in exceedance frequency and duration tracks drought, air temperature, and discharge more clearly than any step change aligned with CCAA implementation. The sections below summarize each report series and independent basin-scale analyses supporting this conclusion.

The monitoring record presented here uses the metrics and units reported in agency documents, and it does not attempt to reanalyze raw temperature time series. However, interpreting these patterns requires clarity about which temperature endpoints are biologically meaningful for persistence in the wild. After synthesizing the monitoring record, we evaluate USFWS's temperature-tolerance framing and propose criteria better aligned with chronic ecological risk.

Data Sources and Approach

We reviewed the Grayling Recovery/Monitoring and CCAA annual report series. Across these reports, the water-temperature monitoring design is consistent: thermographs or loggers at the USGS Wisdom (and later Melrose) gage and a network of mainstem and tributary sites record hourly or sub-hourly temperatures. Early work (1992 onward) relied on thermographs at Wisdom and a handful of upper Big Hole mainstem stations; later work expanded to additional mainstem reaches and key tributaries distributed across CCAA management segments A–E. In all cases the raw data are summarized as daily maximum, minimum, and mean temperatures, seasonal means and maxima, and cumulative hours above 70°F and 77°F.

The monitoring reports and CCAA annual reports all apply a common set of temperature criteria for evaluating grayling habitat. Across the series, field crews summarize each season in terms of mean and maximum temperatures and then tally how often sites exceed two fixed thresholds: 70°F and 77°F. The reports consistently describe ~70°F as a chronic thermal stress threshold for salmonids and 77°F as the "upper incipient lethal temperature" for Arctic grayling based on Lohr et al. (1996). Seasonal summaries therefore focus on seasonal maxima and on cumulative days or hours above 70°F and 77°F at each site. In the review that follows, we use these same thresholds to describe patterns in thermal conditions and to compare results across the different report series.

Early Recovery and Restoration Period

In the early Recovery / Restoration period (1992–early 2000s), the monitoring reports show that summer water temperatures in the upper Big Hole River basin were frequently at or above stress thresholds for Arctic grayling, especially in the mid-basin "warmed reach." During the severe drought of 1992, water temperatures in the upper Big Hole peaked at about 81°F with mean daily temperatures at downstream stations approaching 70°F (Byorth 1993). Even in years with better runoff later in the decade, the mid-basin mainstem sites typically recorded daily maxima in the upper 70s to around 80°F, while headwater sites near Wisdom tended to be cooler, often topping out in the high-60s to low-70s (Magee and Opitz 2000).

By 2003, the network of stream temperature monitoring locations had expanded to roughly a dozen mainstem and tributary loggers, in addition to the Wisdom and Melrose stream gages, and maximum instream temperatures at most stations still occurred in early–mid July, consistent with a basin-wide peak-heat window (Magee and Lamothe 2004). Relative to the 70°F and 77°F thresholds, the early restoration reports show that exceedances were common in hot, dry years and strongly concentrated in the warmed reach. In 2003, water temperatures rose above 70°F at all mainstem stations, and several mid-basin sites exceeded 77°F for many days, with seasonal maxima around 80°F and on the order of 50–70 days above 70°F and 8–17 days above 77°F. Some tributaries were nearly as warm, whereas others remained distinctly cooler and functioned as thermal refugia, rarely if ever exceeding 70°F.

Earlier in the sequence, the 1999 report notes that "lethal" temperatures occurred only at a single station in the warmed reach, with other sites staying below that level, reinforcing that chronic threshold exceedances and near-lethal events were already localized but recurrent features of the system well before the CCAA era (Magee and Opitz 2000).

Mid-2000s to Early 2010s

Across the mid-2000s to early 2010s, the MFWP Arctic Grayling Monitoring Reports for the upper Big Hole show a consistent thermal pattern: relatively cool headwaters and upper mainstem, and a band of chronically warm mid-valley mainstem and tributaries. Seasonal maxima are generally in the mid-60s to low-70s°F at the upper boundary site (Saginaw Bridge) and mainstem Segment A/B sites. By contrast, the warmest conditions occur in mid-basin and lower mainstem sites (for example near the Mudd Creek and Dickie Bridge reaches) and in several low-gradient tributaries, where seasonal maxima repeatedly reach the mid- to upper-70s°F. In individual years, sites such as Governor and Steel creeks reached mid-70s°F, and in 2012 Steel Creek reached 77.2°F, the highest maximum and the only site to exceed the upper incipient lethal temperature that year (Cayer and McCullough 2012).

Looking specifically at threshold exceedances, the MFWP reports document widespread 70°F stress events and intermittent 77°F lethal-range events at a consistent set of "hot" locations. In 2009, roughly two-thirds of monitored mainstem sites and most tributaries exceeded 70°F, and a small subset of mainstem and tributary locations exceeded 77°F; the warmest sites were in the mid-valley mainstem and nearby valley-bottom tributaries (Magee and McCullough 2011). In cooler years such as 2010 and 2011, many sites still experienced multiple days above 70°F, but no locations crossed 77°F. By 2012 and 2013, lethal-range events re-emerged: Steel and Governor creeks and, by 2013, a lower mainstem site near Dickie Bridge all surpassed 77°F with documented hours at or above that level (Cayer and McCullough 2013 and 2014).

Taken together, these reports paint a consistent picture: upper mainstem sites remain comparatively cool, while mid- to lower-valley mainstem reaches and a recurring cluster of low-gradient tributaries are the centers of both 70°F stress and 77°F exceedances.

CCAA Era Annual Reports

The CCAA annual reports document a very similar temperature pattern: cold headwaters and high-elevation tributaries contrasted with consistently warm mid-valley tributaries and mainstem sites. Upper mainstem locations such as Saginaw Bridge (Segment A) and cooler tributaries like LaMarche Creek, upper Steel Creek, Fishtrap Creek, Rock Creek, and portions of French Creek generally have no or very few hours above 70°F and essentially no hours above 77°F. For example, upper mainstem and cold-tributary sites in 2021–2023 typically had seasonal maxima near 70°F and zero hours above either threshold, functioning as reliable thermal refugia.

By contrast, the CCAA reports highlight a recurring cluster of “hot spots” where temperatures routinely exceed the 70°F stress threshold and, in several cases, the 77°F upper incipient lethal threshold for grayling. On the mainstem, mid-valley stations near the Miner Creek and Mudd Creek reaches and the lower-valley site near Dickie Bridge often reach daily maxima of about 73–79°F and accumulate tens to hundreds of hours above 70°F, though hours above 77°F are relatively rare and typically confined to the warmest years. The most thermally stressed sites are low-gradient tributaries and springs in the valley bottom—a recurring hot-spot cluster that includes Big Lake Creek, lower Miner and Steel creeks, Swamp and Plimpton–Howell creeks, Smith Springs, and nearby warm tributaries. A representative example is Big Lake Creek (two monitoring sites). In 2022, Big Lake Creek 1 reached a maximum seasonal temperature of 84.12°F and accumulated 374 hours >70°F and 87 hours >77°F (see Table 9 in MFWP 2022). In 2023, Big Lake Creek 1 again exceeded thresholds, reaching 79.68°F with 283 hours >70°F and 24 hours >77°F (see Table 10 in MFWP 2023). Other streams in this cluster show comparable patterns, with seasonal maxima in the mid-70s°F to lower 80s°F and prolonged periods above stress thresholds.[3]

Together, the CCAA reports portray a stable spatial pattern: persistent warm-water problem areas in specific valley-bottom tributaries and mid-valley mainstem reaches, contrasted with cooler headwaters and spring-fed sites.

Independent Basin-scale Assessments

In addition to the above report series, Sladek (2013) provides a basin-scale check on the same story by analyzing long-term hourly temperature records from the Wisdom and Melrose USGS gages for 1996–2012. H_e calculated annual hours above the salmonid stress threshold (70°F) and a lethal-range threshold (77°F), and then related those exceedances

[3] These site summaries are reported in the annual CCAA temperature monitoring tables, which compile hourly May 1–Oct 1 logger data into seasonal mean/max temperatures and cumulative hours above 70°F and 77°F (MFWP 2023).

to air temperature and discharge rather than to specific management periods. Across the record, Sladek found that temperatures exceeded 70°F in most years and occasionally exceeded 77°F, with no clear downward trend over time and interannual variation largely explained by climate and flow conditions rather than any detectable step change associated with restoration or the CCAA. His conclusion was that more time, more intensive data collection, and formal statistical work would be required before any improvement in water quality could be demonstrated. This independent, gage-based synthesis mirrors the spatial and temporal patterns described in the Early Recovery, MFWP Monitoring, and CCAA annual reports, and it sets up the conclusion that, at least through 2012, there was no indication that CCAA implementation was measurably improving instream temperatures at the basin-scale.

Beyond the annual report series, Vatland's (2015) dissertation provides an independent, basin-wide assessment of summer thermal habitat and climate-driven change for Arctic grayling and non-native salmonids in the upper Big Hole River. Using a combination of airborne thermal infrared imagery, longitudinal temperature profiles, and a network of stationary loggers, he mapped fine-scale summer 2008 temperatures along roughly 100 km of the mainstem, showing pronounced spatial heterogeneity with relatively cool headwater and spring-fed segments and substantially warmer mid-valley and lower mainstem reaches. H_e then coupled site-specific air–water temperature regressions with regional climate model projections to simulate weekly mean daily maximum water temperatures from the 1980s through the 2060s, evaluating exceedance of chronic (21.1°C) and acute (25°C) thermal thresholds across the stream network. The simulations indicated significant basin-wide warming over coming decades, with sharp increases in the frequency and spatial extent of threshold exceedances in already-warm reaches and the development of contiguous warm segments that could function as seasonal thermal barriers to grayling movement, even as some high-elevation and groundwater-influenced sections were projected to remain below thresholds and serve as refugia. Overall, Vatland's work reinforces the picture from the monitoring reports of long-standing hot reaches and cold-water refuges, while projecting that climate change alone is likely to intensify existing thermal stress patterns rather than eliminate them.

Implications for the CCAA

USFWS has argued that Vatland's (2015) projections are conservative because they are based on a static landscape and do not incorporate conservation measures implemented after 2010, including later CCAA actions. That critique is fair as far as it goes—Vatland did not attempt to model the effects of subsequent restoration or flow-management changes. However, the CCAA annual reports and associated monitoring documents covering 2006–2023 do not show a clear, basin-wide reduction in the frequency or duration of temperature exceedances consistent with the CCAA measurably offsetting the climate-driven warming Vatland projected. Instead, high summer temperatures and recurrent exceedances of 70°F and 77°F persist at many of the same mainstem and tributary "hot spots" identified in Vatland's work and in earlier monitoring. In that sense, the empirical record to date does not support the theory that CCAA implementation has measurably mitigated the thermal effects of a warming climate in the upper Big Hole.

Temperature Trends Conclusion

Taken together, the early Recovery/Restoration reports, the MFWP Arctic Grayling Monitoring Reports, and the CCAA annual reports do not provide evidence that implementation of the CCAA has measurably improved instream water temperatures in the upper Big Hole River basin. All three eras show the same basic pattern: relatively cool headwater/mainstem sites and a recurring set of mid-valley mainstem and low-gradient tributary “hot spots” that frequently exceed the 70°F stress threshold and, in some years, the 77°F lethal threshold. Exceedances occur before, during, and well into the CCAA period, and year-to-year differences in frequency and duration track drought, snowpack, and flow conditions much more clearly than any step-change associated with CCAA implementation.

While the CCAA reports emphasize projects (flow leasing, irrigation upgrades, riparian work) that are intended to reduce high temperatures and may well provide localized benefits, the monitoring record as summarized in these documents does not show a consistent, basin-wide decline in the occurrence or duration of threshold exceedances. In short, the reports indicate ongoing, spatially persistent thermal stress with substantial interannual variability, rather than a clear pattern of improved instream temperatures attributable to the CCAA. Differences in site coverage, metric definitions (days vs hours, temperature reported in °F vs °C), and climate conditions among eras limit strict before–after statistical comparisons, but the overall spatial pattern and persistence of threshold exceedances are similar across all three report series.

How Temperature Risk is Defined

USFWS’s ESA risk analyses began to shift toward higher, more permissive temperature benchmarks in the 2010s. Earlier 12-month findings anchored chronic risk to a conservative 20°C (68°F; 7-day average of daily maxima), consistent with the Hubert et al. (1985) habitat suitability model that treats grayling habitat as unsuitable above ~20°C. In the 2020 “Not Warranted” finding, however, USFWS shifted to treating 21.1°C as the chronic stress benchmark and 25–26.9°C (UILT/CTM) as “ecological thresholds” for survival, and evaluates risk using weekly means and total hours per year above those higher numbers. In other words, while the monitoring reports reviewed in the previous section flag exceedances of 70°F and 77°F as biologically important events, the later federal analyses reframe grayling as more tolerant by raising the chronic benchmark and elevating acute laboratory endpoints (UILT/CTM) into de facto management thresholds.

In this section, we recommend using the conservative chronic threshold of 20°C for evaluating threats to Arctic grayling. We explain why the two acute lethal lab metrics—upper incipient lethal temperature (UILT) and critical thermal maximum (CTM)—do not work as criteria for an ESA risk assessment. We also explain why metrics tied to physiological response and sub-lethal effects are more appropriate than acute laboratory endpoints that only tell you when fish die.

Why Endpoint Choice Matters

Researchers use UILT and CTM to measure short-term mortality under controlled conditions (constant high temperature for UILT; rapid ramping to loss-of-equilibrium for CTM). They do not represent temperatures at which Arctic grayling can maintain ecological performance—feed, grow, and reproduce across available habitat. They also do not mirror natural conditions where factors such as oxygen levels, food availability, disease resistance, predation, and competition influence the fish's ability to resist high temperatures.

Why UILT/CTM are unsuitable as risk assessment criteria:

- **Endpoint mismatch:** These tests measure short-term survival, but persistence depends on long-term costs (feeding, growth, disease, movement, reproduction).
- **Method sensitivity:** Acclimation temperature, ramp rate, and body size can shift UILT/CTM results.
- **Ramp-rate artifacts:** Rapid temperature ramping can inflate apparent tolerance because internal body temperature and physiological strain lag behind external water temperature.

In the CTM tests cited in USFWS's 2020 finding, ramp rates were ~14–34 times faster than typical river warming. Feldmeth and Eriksen (1978) raised the temperature in test tanks at a rate of 0.16-0.19°C per minute and Lohr et al. (1996) raised it at a rate of 0.4°C per minute. By contrast, Big Hole River summer warming typically occurs on the order of ~0.7°C/hour (median hourly warming rate during June–August, 2012–2025, from the USGS-06024450 gage record). When water warms quickly, a fish's body temperature lags behind water temperature, so loss-of-equilibrium is reached before tissues fully equilibrate or cumulative damage accrues. In fact, the faster rate of change in CTM tests was designed to prevent acclimation during the test (Desforges et al. 2023).

USFWS (2020) argues that UILT and CTM tests are "conservative" for wild fish because those lab methods don't incorporate diel (daily) temperature cycling and exposure duration. They then make the more specific claim that fish experiencing daily fluctuations can survive higher peak temperatures than fish held at constant (UILT) or steadily increasing (CTM) temperatures. In support of these claims, USFWS provides two references: Johnstone and Rahel 2003, and Dickerson and Vinyard 1999. The papers cited by USFWS, as well as others discussed below, do not support the concept that laboratory-derived UILT/CTM should be viewed as "conservative estimates when analyzing potential effects of water temperature."

Johnstone and Rahel (2003) observed markedly different survival/behavioral responses in cutthroat trout exposed to constant vs cycling temperatures. Under a 7-day constant regime their upper temperature threshold was 24.2°C and fish survived a 7-day diel cycle of

16–26°C. But crucially, they also observed declines in feeding and activity during the diel cycle treatment and state that long-term exposure would be detrimental. They also stated that mortality occurred when cycles were increased to 18–28°C. These were short-term tests, and the Big Hole River can exceed the chronic threshold from May through September (USGS-06024450 record).

Dickerson and Vinyard (1999) also note that cycling can reduce mortality relative to constant exposure, but the authors' interpretation is not "higher tolerance" in the way that USFWS has interpreted it. They explicitly conclude that acclimation to fluctuating temperatures did not increase tolerance to chronic temperatures; rather fish likely persist inactive during the above-limit period and resume activity when temperatures drop. They also warn that physiological survival can still mean "ecological death" because fish would be "incapable of finding food or avoiding predators under field conditions."

Research has shown that night-time recovery does not necessarily mitigate daily temperature peaks. Salmonid studies using diel-fluctuating regimes show that daily peaks carry disproportionate costs: juvenile salmonids exposed for ~30 days to matched mean temperatures grew nearly twice as fast in constant regimes compared to fluctuating regimes, and short daily peaks near 27°C caused significant mortality despite cool nights (Geist et al. 2010). Classic rainbow trout work (Hokanson et al. 1977) shows that once daily fluctuations extend above the growth optimum, growth is worse—and mortality higher—than at the same mean under constant conditions. In other words, surviving a peak is not the same as being protected from chronic, cumulative harm.

Chronic Stress Thresholds

Instead of focusing on the lethal temperatures derived from laboratory tests, we recommend USFWS use a chronic stress threshold temperature. Recent grayling-specific research supports a conservative chronic threshold well below UILT/CTM. In a modern ~144-day experiment with Big Hole River basin origin juveniles, growth peaked around 18°C, survival was stable up to 20°C and declined sharply at 22°C, and no survival occurred ≥24°C (Carrillo-Longoria et al. 2023). Indicators of oxidative stress also increased above 16°C. These results indicate that sub-lethal harm and survival failure occur well below UILT/CTM, confirming the 20°C chronic threshold originally proposed by Hubert et al. (1985) and used in two USFWS 12-month findings (75 FR 54708, 79 FR 49384). By contrast, USFWS's 2020 finding shifted to 21.1°C as the temperature at which Arctic grayling begins to experience physiological stress and treats 25-26.9°C as ecological thresholds for survival. The rationale presented for the shift to 21.1°C appears to rely heavily on a brook trout ecological limits study (Chadwick et al. 2015) rather than grayling-specific performance data.

A standard practice in western guidance is to pair an average-of-peaks metric with peak-aware indicators. USFWS's 12-month finding discusses the temperature evaluation in Vatland (2015) that used a weekly mean of daily maxima at 21.1°C and 25°C. In response, USFWS states that actions implemented under the CCAA mitigate the threat identified in Vatland (2015). From here USFWS shifts to an evaluation of the number of hours per year that exceeded these same temperature thresholds and demonstrates that the annual hours have declined. The analytical focus on hours per year, however, does not capture the peak magnitude and time-above-threshold that experiments identify as the primary drivers of risk (hours per year can decline even if peak magnitude and time-above-threshold are still biologically harmful).

UILT/CTM do not work as ESA ‘thresholds’ because they ignore the sub-lethal, cumulative, and peak-driven processes that determine whether grayling populations persist. The best available grayling science indicates that sub-lethal harm begins well below lethal limits: growth and performance at 18°C, physiological stress increases above ~16°C, and survival declines as exposures approach the low-20s—especially when daily peaks recur (Carrillo-Longoria et al. 2023). In rivers, grayling experience diel-cycling heat, not steady laboratory baths; in that setting, peak intensity and time-above-threshold drive cumulative damage that night-time cooling does not fully erase.

Recommended Temperature Criteria

We recommend an assessment that emphasizes peak-aware, sub-lethal criteria anchored to physiological response, growth, migration, and disease resistance—so that the Upper Missouri River Arctic grayling is protected before temperatures reach the lethal cliff. In addition to using the scientifically supported 20°C chronic protection threshold for Big Hole River Arctic grayling, USFWS should use a 7-day average daily maximum (7DADM) metric as the primary chronic indicator and the number of days > 20°C and the longest run of consecutive days above the same threshold as secondary indicators.

Cross-chapter Synthesis

Taken together, the flow and temperature chapters describe a chronic summer/early-fall habitat squeeze—low flows coincident with elevated temperatures—that can reduce the amount of suitable rearing habitat. This pattern aligns with the demographic and genetic signals seen in catch-per-unit-effort (CPUE) and effective number of breeders (N_b): periodic recruitment signals without sustained gains in older fish.

Across the indicators reviewed, the monitoring record does not show clear evidence of improved viability during the CCAA era. Fall CPUE provides no evidence of increasing abundance of older fish: Tributaries Age-1+ declined, and Mainstem Age-1+ showed no positive trend. Meanwhile, episodic young-of-year (YOY) increases have not carried forward into higher Age-1+ CPUE, suggesting a bottleneck between early recruitment and recruitment to older age classes. N_b is broadly consistent with this interpretation because it reflects successful breeding and early recruitment, not the abundance of older fish.

We did not directly test juvenile survival mechanisms, but reduced juvenile habitat availability during warm, low-flow periods is a plausible limiting pathway. Alternative hypotheses (e.g., limits on Age-1+ habitat capacity or food resources) would produce a similar “recruitment without retention” pattern and warrant targeted evaluation.

The following section explains how this integrated evidence maps to ESA listing considerations.

Overall Conclusions & ESA Listing Recommendation

The Upper Missouri River DPS encompasses multiple populations across several river systems and isolated mountain-lakes. Recent synthesis work describes 19 extant populations across the broader conservation portfolio, including historically occurring, naturally established populations in the upper Big Hole and Centennial valleys and 12 introduced mountain-lake populations (Montana Arctic Grayling Workgroup 2022, Kreiner 2025).

This report does not attempt to resolve debates about the conservation value of introduced lake populations or whether they should be considered a substitute for historically occurring, naturally established populations. Instead, we emphasize the Big Hole River because it is central to DPS representation of fluvial life history and because its monitoring record provides direct evidence of how threats translate into demographic outcomes.

Across independent indicators, the evidence does not support a finding of improved viability. Abundance metrics show no sustained increase in older fish, and recruitment signals have not translated into sustained gains. Habitat analyses indicate that summer/fall streamflow deficits and elevated summer/fall temperatures remain chronic, reinforcing a mechanistic link between the dominant threats (low flow and high temperature) and the observed demographic patterns (YOY pulses have not translated to increased abundance of older age classes). Because the Big Hole is the DPS's core fluvial population, the absence of measurable improvement here has major implications for DPS-wide viability and redundancy.

Accordingly, we conclude that the Upper Missouri River DPS warrants listing under the Endangered Species Act. In our assessment, ESA listing is justified based on the continued exposure to key threats, the lack of demonstrated improvement in viability, and the high consequence of further losses in the remaining core fluvial population. Effective population size (N_e) also supports this conclusion: current monitoring indicates N_e remains below widely cited long-term benchmarks for maintaining adaptive potential, and therefore does not demonstrate secure long-term genetic resilience in the foreseeable future—especially under persistent habitat stressors.

This recommendation is consistent with the administrative and judicial record. USFWS's position on listing has shifted over time, including a 2007 determination that the population was not a listable entity, a 2010 determination that the DPS was listable and warranted listing but was precluded, and subsequent "not warranted" findings that have been vacated in federal court. In our view, the best available science summarized here supports ESA listing for the Upper Missouri River DPS.

Mapping the Evidence to ESA section 4(a)(1)

Key listing factors supported by this report include:

Factor A (habitat destruction/modification): chronic summer/fall low flows and summer/fall temperature exceedances reduce available habitat, increase physiological stress, and constrain growth and survival.

Factor D (inadequacy of regulatory mechanisms): the voluntary CCAA and associated measures have not produced measurable, basin-scale improvements in summer/fall habitat conditions relative to the biological needs of the species.

Factor E (other natural/manmade factors—climate change): ongoing climatic warming is increasing thermal stress and amplifying the effects of low summer/fall flows, reducing the margin for persistence in a small, fragmented DPS. Persistent summer/fall temperature exceedances in the core fluvial population support climate change as a continuing threat.

References

Allendorf, F.W., Hössjer, O., & Ryman, N. (2024). What does effective population size tell us about loss of allelic variation? Evolutionary Applications, e13733.

Big Hole River Watershed Committee. 1997. Big Hole River Drought Management Plan. Adopted 1997. Amended 1999, 2000, 2002, 2004, 2005, 2007, 2008.

Byorth, P.A. 1991. Population surveys and analysis of fall and winter movements of Arctic grayling in the Big Hole River. 1991 Annual Report. Submitted to: Fluvial Arctic Grayling Workgroup. Montana Fish, Wildlife & Parks, Bozeman, MT.

Byorth, P.A. (1993). Big Hole River Arctic Grayling Restoration Project Annual Report: 1992. Unpublished report submitted to: the Fluvial Arctic Grayling Workgroup and partners.

Carrillo-Longoria, J.-A., Gaylord, G., Andrews, L., & Powell, M. (2023). Effect of temperature on growth, survival, and chronic stress responses of Arctic Grayling juveniles. Transactions of the American Fisheries Society, 153(1), 3–22.

Cayer, E. and A. McCullough. (2012). Arctic grayling monitoring report. Montana Fish, Wildlife, and Parks, Dillon Field Office, 32 pp.

Cayer, E. and A. McCullough. (2013). Arctic grayling monitoring report. Montana Fish, Wildlife, and Parks, Dillon Field Office, 26 pp.

Center for Biological Diversity v. Zinke, 900 F.3d 1053 (9th Cir. 2018).

Center for Biological Diversity v. U.S. Fish & Wildlife Service, Case 2:23-cv-00002-DLC, Document 52 (D. Mont. Aug. 6, 2024).

Chadwick Jr, J.G., Nislow, K.H. and McCormick, S.D. (2015). Thermal onset of cellular and endocrine stress responses correspond to ecological limits in brook trout, an iconic cold-water fish. Conservation Physiology, 3(1).

Desforges, J. E., Birnie-Gauvin, K., Jutfelt, F., Gilmour, K. M., Eliason, E. J., Dressler, T. L., McKenzie, D. J., Bates, A. E., Lawrence, M. J., Fangue, N., & Cooke, S. J. (2023). The ecological relevance of critical thermal maxima methodology for fishes. Journal of Fish Biology, 102(5), 1000–1016.

Dickerson, B. R., & Vinyard, G. L. (1999). Effects of high chronic temperatures and diel temperature cycles on the survival and growth of Lahontan cutthroat trout. Transactions of the American Fisheries Society, 128(3), 516–521.

Feldmeth, C. R., & Eriksen, C. H. (1978). A hypothesis to explain the distribution of native trout in a drainage of Montana’s Big Hole River. Verhandlungen der Internationalen Vereinigung für Theoretische und Angewandte Limnologie, 20, 2040–2044.

Frankham, R., Bradshaw, C.J.A., & Brook, B.W. (2014). Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. Biological Conservation, 170, 56–63.

Geist, D. R., Abernethy, C. S., Hand, K. D., Cullinan, V. I., Chandler, J. A., & Groves, P. A. (2010). Survival, development, and growth of juvenile Chinook salmon: Temperature–oxygen ratio. Transactions of the American Fisheries Society, 139(2), 601–612.

Hamed, K.H., & Rao, A.R. (1998). A modified Mann–Kendall trend test for autocorrelated data. Journal of Hydrology 204:182–196.

Helsel, D.R., & Hirsch, R.M. (2002). Statistical Methods in Water Resources. U.S. Geological Survey, Techniques of Water-Resources Investigations, Book 4, Chapter A3.

Hokanson, K.E., Kleiner, C.F. and Thorslund, T.W. (1977). Effects of constant temperatures and diel temperature fluctuations on specific growth and mortality rates and yield of juvenile rainbow trout, Salmo gairdneri. Journal of the Fisheries Board of Canada, 34(5), pp.639–648.

Hubert, W.A., R.S. Helzner, L.A. Lee and P.C. Nelson. (1985). Habitat suitability index models and instream suitability curves: Arctic grayling riverine populations. US Fish and Wildlife Service Biological Report 82(10.110). 34 pp.

Jamieson, I.G., & Allendorf, F.W. (2012). How does the 50/500 rule apply to minimum viable populations? Trends in Ecology & Evolution 27(10):578–584.

Johnstone, H.C., F.J. Rahel. *2003). Assessing Temperature Tolerance of Bonneville Cutthroat Trout Based on Constant and Cycling Thermal Regimes, Transactions of the American Fisheries Society, Volume 132, Issue 1. Pages 92–99.

Kovach, R.P., Whiteley, A.R., Painter, S., Lodmell, A., & Leary, R.F. (2019). Multi-decade genetic monitoring of Arctic grayling in Montana. Montana Conservation Genetics Laboratory Report 2019/1.

Kovach, R.P., Whiteley, A.R., Jaeger, M.E., Painter, S., Lodmell, A., & Leary, R.F. (2020). Genetic monitoring informs conservation of Arctic grayling in Montana. Canadian Journal of Fisheries and Aquatic Sciences.

Kovach, R.P., Painter, S., & Lodmell, A. (2025). Arctic grayling genetic monitoring: Big Hole River and Upper Red Rock Creek (annual monitoring report; January 2, 2025). University of Montana Conservation Genetics Laboratory, prepared for Montana Fish, Wildlife & Parks.

Kreiner, J.M., Kovach, R.P., Painter, S., and collaborators. (2025). Genetic status and ancestry of Arctic grayling lake populations in the Upper Missouri River basin, Montana.

Lohr, S.C., P.A. Byorth, C.M. Kaya, and W.P. Dwyer. (1996). High-temperature tolerances of fluvial Arctic grayling and comparisons with summer river temperatures of the Big Hole River, Montana. Transactions of the American Fisheries Society 125:933–939.

Luikart, G., F.W. Allendorf, J.-M. Cornuet, and W.B. Sherwin. (1998). Distortion of allele frequency distributions provides a test for recent population bottlenecks. Journal of Heredity 89(3): 238–247. DOI: 10.1093/jhered/89.3.238.

Magee, J., S.T. Opitz. (2000). Montana Fluvial Arctic Grayling Recovery Project: Annual Monitoring Report 1999. Submitted to the Fluvial Arctic Grayling Workgroup and partners.

Magee, J., P. Lamothe. (2004). Fluvial Arctic Grayling Monitoring Report 2003. Submitted to the Fluvial Arctic Grayling Workgroup and partners.

Magee, J., A. McCullough. (2011). Arctic Grayling Recovery Program, Montana Arctic Grayling Monitoring Report 2009. Submitted to the Fluvial Arctic Grayling Workgroup and partners.

Magee, J., A. McCullough, and E. Cayer. (2012). Arctic Grayling Monitoring Report 2010. Montana Fish, Wildlife & Parks, Dillon Field Office. Submitted to the Fluvial Arctic Grayling Workgroup and partners.

Montana Arctic Grayling Workgroup. (2022). Upper Missouri River Arctic grayling conservation strategy.

Montana Fish, Wildlife & Parks. (2022). Montana streams: base hydrography for the Montana Fisheries Information System.

Montana Fish, Wildlife & Parks. (2022). Candidate conservation agreement with assurances for fluvial Arctic grayling in the upper Big Hole River: 2022 annual report.

Montana Fish, Wildlife & Parks. (2023). Candidate conservation agreement with assurances for fluvial Arctic grayling in the upper Big Hole River: 2023 annual report.

Montana Fish, Wildlife & Parks. (2025). FishMT survey and inventory data, Big Hole River basin. Accessed November 7, 2025. Retrieved from https://myfwp.mt.gov/fishMT/reports/surveyreport

National Research Council. (2011). Reference Manual on Scientific Evidence. 3rd ed. National Academies Press.

Nei, M., T. Maruyama, and R. Chakraborty. (1975). The bottleneck effect and genetic variability in populations. Evolution 29(1): 1–10. DOI: 10.1111/j.1558-5646.1975.tb00807.x.

Sladek, H. (2013). Trend analysis of water temperature relative to air temperatures and flow in the Big Hole River. Report to the Big Hole Watershed Committee and Montana Department of Water Quality, 42 pp.

Streu, J. M. 1990. Select aspects of the life history and ecology of the Montana Arctic grayling (*Thymallus arcticus montanus* [Milner]) in the upper Big Hole River drainage, Montana: June 15 to August 31, 1990. Report to: Montana Natural Heritage Program, Beaverhead National Forest, and Montana Dept. of Fish, Wildl., and Parks, Bozeman.

U.S. Fish and Wildlife Service. (2020). Revised 12-month finding on a petition to list the Upper Missouri River Distinct Population Segment of Arctic grayling under the Endangered Species Act (and supporting documents).

U.S. Fish and Wildlife Service. (2006). Candidate Conservation Agreement with Assurances for Fluvial Arctic Grayling in the Upper Big Hole River. Between Montana Department of Fish, Wildlife and Parks and U.S. Fish and Wildlife Service In cooperation with Montana Department of Natural Resources and Conservation and USDA Natural Resources Conservation Service.

U.S. Fish and Wildlife Service. (1982). Endangered and threatened wildlife and plants; review of vertebrate wildlife for listing as endangered or threatened species. Federal Register 47:58454 (December 30, 1982).

U.S. Fish and Wildlife Service. (1994). Endangered and threatened wildlife and plants; finding on a petition to list the fluvial Arctic grayling as endangered. Federal Register 59:37738 (July 25, 1994).

U.S. Fish and Wildlife Service. (2004). Endangered and Threatened Wildlife and Plants; Review of Species That Are Candidates or Proposed for Listing as Endangered or Threatened; Annual Notice of Findings on Resubmitted Petitions; Annual Description of Progress on Listing Actions. Federal Register 69:24876 (May 4, 2004).

U.S. Fish and Wildlife Service. (2010). Endangered and Threatened Wildlife and Plants; Revised 12-Month Finding to List the Upper Missouri River Distinct Population Segment of Arctic Grayling as Endangered or Threatened. Federal Register 75:54708 (September 8, 2010).

Vatland, S.J. (2015). Effects of stream temperature and climate change on fluvial Arctic grayling and non-native salmonids in the upper Big Hole River, Montana. Dissertation, Montana State University, 128 pp.

Wasserstein, R.L., & Lazar, N.A. (2016). The ASA Statement on p-values: context, process, and purpose. The American Statistician 70(2):129–133.