WEST VIRGINIA RIVERS

April 15, 2022

WV Department of Environmental Protection
601 57th Street South East
Charleston, WV

Attn: Brian Bridgewater

Re: 2022 Triennial Review of Water Quality Standards

Mr. Bridgewater:

West Virginia Rivers Coalition, on behalf of our members and the organizations signed below, respectfully submits the following comments on the WV Department of Environmental Protection's (WVDEP) on Rules Governing Water Quality Standards (47CSR2) as part of the 2022 Triennial Review.

Human Health Criteria

In 2015, EPA updated its National Recommended Water Quality Criteria for human health for 94 chemical pollutants to reflect the latest scientific information and EPA policies, including updated fish consumption rate, body weight, drinking water intake, health toxicity values, bioaccumulation factors, and relative source contributions. During the 2019 triennial review, WVDEP chose to only update the recommended criteria for pollutants which the WVDEP already had existing standards.

Our concern is that our citizens remain vulnerable to health risks posed by toxins and carcinogens in EPA's 2015 recommended criteria updates that remain unregulated in West Virginia. Through our participation in WVDEP's Human Health Criteria Workgroup, WVDEP stated that it's intention was to first address the recommended updates to criteria that WV currently had existing standards for, and then it would consider the adoption of the remaining recommended updates of which WV has yet to establish standards. Now is the time to consider adopting the remaining 2015 recommended criteria; there should be no further delay.

DRAFT

We request that WVDEP adopt the remaining human health criteria included in EPA's 2015 recommended updates so that West Virginia has more comprehensive standards in place to adequately protect public health.

B2 Trout Waters

The current B2 trout waters list in Appendix A of the state's water quality standards, underrepresents the state's exceptional trout resources, and has not been updated for many years. We are aware that WVDEP is currently assessing the state's trout stream list for input into EPA's database. This triennial review is the right time for WVDEP to undertake a thorough and comprehensive assessment of the state's trout waters and make sure that our water quality standards are fully protective of all trout waters.

In partnership with Trout Unlimited, WV Rivers embarked on a process to produce a wild trout atlas that compiled WVDEP and WVDNR trout stream lists and identified gaps in B2 trout water designation. We also formed the WV Wild Trout Collaborative, a workgroup composed of conservation organizations and state and federal agencies tasked with managing trout resources in WV. Through this process, we have so far identified 53 streams containing trout DNA that are not included on the WVDEP or WVDNR trout stream list. This is one example of the inadequacies of the current trout list that illustrates the need to reassess how trout streams are defined and identified in the state, and streamline a process to add additional trout streams to the list.

WVDEP must ensure that all trout waters in the state have adequate protections. We understand this process deserves careful scientific review and urge the WVDEP to immediately begin research and consultation with stakeholders and the scientific community.

We request the WVDEP form a workgroup to reassess the state's B2 trout waters in Appendix A, develop a process to update the state's trout stream list, and recommend revisions to 47CSR2 related to trout waters as part of this triennial review. We recommend this workgroup be comprised of representatives from the WV Wild Trout Collaborative, as well as other scientists with specialized expertise in trout.

DRAFT

Conductivity Standard

WVDEP needs to implement a water quality standard based on ionic toxicity from surface mining operations. The scientific literature has demonstrated for more than 10 years that such a standard is necessary to protect aquatic life. Indeed in 2011, EPA scientists Susan Cormier and Glenn Suter established a benchmark for conductivity to be used for just such a purpose. (EPA, 2011). The methodology was peer-reviewed and published in 2013. (Cormier and Suter, 2013).

The scientific justification has only become more robust since then. Appendix A includes a literature review that summarizes the support for a conductivity threshold of 300 µS/cm for protection of aquatic life from ionic pollution from surface mines. The causal link between highly conductive ionic runoff and a violation of the existing water quality standard for aquatic life has been recognized by federal judges in West Virginia and the Fourth Circuit Court of Appeals, which are also summarized in Appendix A.

We request the WVDEP implement a water quality standard for ionic pollution from surface mines at a threshold of 300µS/cm. Failure to do so would be a gross disregard of the science and the law.

Selenium Standard

Since the implementation of fish-tissue based selenium standards, coal mining operations across the state have abandoned planned, or even active, selenium treatment systems because of the reduced stringency of that standard. Though the WVDEP guidance was written in 2016 and updated in 2018, it fails to take account of the findings of a 2015 paper by USGS scientist Nathaniel Hitt.

In that paper, Dr. Hitt raised serious concerns about the implementation of the tissue standard without specific limitations on sample sizes that would be used to assess attainment. Dr. Hitt cautioned that sample sizes of at least 8 fish would be required to detect increases of 1.2 mg/kg Se above the EPA threshold with 80% confidence (at a type-1 error threshold of 0.20). Further, samples of fewer than 5 fish would be unable to detect increased mean Se concentrations of 1.7 mg/kg with even an 80% confidence.

DRAFT

Currently WV's guidance on fish tissue does not necessitate sample sizes of greater than even 5 fish. *See* WV Selenium Aquatic Life Standard Implementation at 4-5 (rev.4 2018). For example, a composite sample with 3 fish of the same species and one of a different minnow group could be combined for analysis resulting in a 4-fish sample. *Id.* Given the strong motivation for companies to avoid costly selenium treatment, the standard should be revised to eliminate Type-II errors (in other words, errors that fail to recognize selenium concentrations above the threshold.)

We request that a requirement for an 8-fish minimum composite sample be adopted to determine compliance with the fish-tissue based selenium standard.

Dr. Hitt's paper is attached in Appendix B for reference.

Transitioning from fecal coliform to E. coli for bacteria standard

In 2012, EPA recommended using Escherichia coli (E. coli) as an indicator of fecal contamination for fresh water. The neighboring states of Virginia, Pennsylvania, Ohio and Kentucky all use E. coli to protect primary contact recreational uses in surface waters. For the past 10 years, WV has resisted making the transition from fecal coliform to E. coli. In 2016, the transition was attempted but later abandoned. Other states, such as Colorado, adopted dual fecal coliform and E. coli criteria in anticipation of the transition.

E. coli reflects stronger scientific links between potential sources of bacteria and likelihood of human illness following contact with water contaminated by fecal material. The most recently recommended indicators represent types of bacteria that are more specifically linked to pollution sources related to fecal material from warm blooded animals. This means that measurement of E. coli provides a more defensible link to pollution sources like malfunctioning wastewater treatment plants or failing home septic systems. Because of this and the more defensible correlation with human illness, E. coli is a better indicator and a more modern and protective standard for water quality criteria.

We request that DEP develop a process to make the transition from fecal coliform to E.coli as the contact recreation standard.

Thank you for your careful consideration of these comments. We look forward to discussing our recommendations at the public meeting on June 28, 2022.

Signed,

Angie Rosser
West Virginia Rivers Coalition

DRAFT

Appendix A Scientific and Judicial Justification for Conductivity Standard

Literature Review

Surface coal mining in the Appalachian region of relevance here requires disturbing and removing overburden rock layers to access coal-rich deposits, which exposes pyrite-rich spoils to rainwater, accelerating the dissolution of naturally occurring metals and other cations (e.g., HCO3, SO4, Ca, Mg, Fe, Mn, Se, K, NO3/NO2) in the process. In fact, the accelerated chemical weathering measured below mountaintop mines are the highest ever recorded (Ross et al. 2018). The ionic soup and elevated conductance produced by the above processes is a characteristic signal of surface coal mining in southern West Virginia, southwestern Virginia, and eastern Kentucky (Pond et al. 2008, US EPA 2011, Lindberg et al. 2011, Griffith et al. 2012, Cormier et al. 2013a, Kunz et al., 2013; Merriam et al. 2015).

Early Studies on Water Quality Impacts of Surface Mining. Many studies since 2005 have reported that streams below surface mines and valley fills which exhibit elevated conductivity (e.g., >300 µS/cm, but often >> 1000 µS/cm) also show signs of biological impairment, and these studies have implicated ions contributing to stream conductivity as the most likely drivers (Cormier et al. 2013a, b). These articles cumulatively have more than fifty authors and have been peer-reviewed by dozens of eminent scientists.

The first article documenting a relationship between loss of biota and elevated conductivity in alkaline streams of Appalachia was Hartman et al.'s 2005 paper in *Hydrobiologia*. Using a paired study design, they showed substantial degradation of the benthic macroinvertebrate community associated with elevated conductivity. Pond et al.'s 2008 peer-reviewed study in the *Journal of the North American Benthological Society* stated in the abstract:

> We characterized macroinvertebrate communities from riffles in 37 small West Virginia streams (10 unmined and 27 mined sites with valley fills) sampled in the spring index period (March–May) and compared the assessment results using family- and genus-level taxonomic data. Specific conductance was used to categorize levels of mining disturbance in mined watersheds as low (500 µS/cm), medium (500–1000 µS/cm), or high (1000 µS/cm). Four lines of evidence indicate that mining activities impair biological condition of streams: shift in species assemblages, loss of Ephemeroptera taxa, changes in individual metrics and indices, and differences in water chemistry.

Pond et al. (2008) showed that macroinvertebrates associated with low-conductivity, low-sulfate reference streams decline sharply with increasing conductivity associated with alkaline mine drainage (i.e., sites with elevated sulfate, low chloride, and basic pH). Pond et al. (2008) further found that "[o]ur results confirm that MTM impact to aquatic life is strongly correlated with ionic strength in the Central Appalachians, but habitat quality did explain some variance in MMIs and other metrics." *Id.* at 725.

In 2010, Pond published a peer-reviewed paper in *Hydrobiologia* which found that in eastern Kentucky "[m]ean mayfly richness and relative abundance were significantly higher at REF [reference] sites compared to all other categories; MINED sites had significantly lower metric values compared to RESID [residential] and MINED/RESID sites." He further stated that "[a]nalyses from WV mining areas . . . indicated that the decline of mayflies from mountaintop mining correlates most strongly to specific conductance." *Id*. at lines 603-607. Thus, Pond found that mayflies declined or were eliminated from mined areas and that the abundance of mayflies was more closely related to conductivity than to habitat.

In linking surface mining to biota, Gerritsen et al. (2010) demonstrated that surface coal mining had a distinct biotic signal distinguishing chemical stress from habitat alteration and altered food resources in particular (temperature was not a significant factor in this case). This study illustrates that it is possible to decompose lowered multi-metric benthic condition scores (e.g., the West Virginia Stream Condition Index—WVSCI; Virginia Stream Condition Index—VSCI) scores using assemblage composition to distinguish the characteristic biotic signals of surface mining.

In a geographic study of mining extent, Petty et al. (2010) found that streams in catchments with very low (i.e., 1-5%) mining intensity were nonetheless biologically degraded, and these streams consistently had elevated conductivity and sulfate. Similarly, Merriam et al. (2010) found that in streams degraded by moderate amounts of either human development or mining, mining further degraded developed streams, highlighting the distinctively harmful effect of mining impacts.

Palmer et al. (2010) published a peer-reviewed study in *Science*, a premier scientific journal. They found that as mining increased, conductivity and sulfate increased, and biological condition (as measured by WVSCI) declined, including a decline in mayflies.

Merriam et al. published a peer-reviewed paper in early 2011 in the *Journal of the North American Benthological Society* on the effects of mining and residential development in Central Appalachia. The paper found that "mining (% of total subwatershed area) caused acute changes in water chemistry," that sites affected by mining and development "had lower Ephemeroptera, Plecoptera, Trichoptera richness than sites affected by either stressor alone," and that the biological impairment threshold was breached when mining activities covered about 25% of the cumulative subwatershed area. (abstract & p. 411). The study's authors "observed biological impairment when conductance reached 250 µS/cm." (pp. 413-14). More recently, Merriam et al.

DRAFT

(2015) showed consistently positive effects of surface mining on conductivity and consistently negative effects on biotic integrity throughout southern West Virginia. These effects, though variable in degree across local watersheds, were exacerbated by residential development and the presence of underground mines.

The EPA Benchmark Study. In 2011, EPA scientists issued a report called “A Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams.” The Benchmark was authored by scientists like Cormier and Suter, who had published important papers in the area of ecological causation. (p. ix). Pond was also a contributor to the report. (p. x). Before publication, the Benchmark was reviewed by a scientific advisory board, which itself was composed of top scientists who possessed expertise in the area. (pp. xi-xii). The Benchmark used EPA’s standard method for deriving water-quality criteria to derive a conductivity benchmark of 300 µS/cm. (p. xiv). This method was itself later published and subjected to independent peer review by the authors (Cormier et al., 2013c, Cormier and Suter, 2013). Under that method, EPA sets the benchmark at the level needed to protect 95% of macroinvertebrate taxa. (p. xiv). Figure 8 in the benchmark graphs the species sensitivity distribution and shows that extirpation increases as conductivity increases. (p. 18). Five percent of taxa are lost when conductivity rises to 295 µS/cm, over 50% are lost at 2000 µS/cm, and close to 60% are lost at 3000 µS/cm. (p. 18).

EPA conducted a detailed causal assessment and concluded that there is a causal relationship between conductivity and stream impairment in West Virginia. (pp. 40, A-40 (“This causal assessment presents clear evidence that the deleterious effects to benthic invertebrates are caused by, not just associated with, the ionic strength of the water. . . . When [other potential] causes are absent or removed, a relationship between conductivity and ephemeropteran [, i.e. mayfly,] richness is still evident.”) Again, the authors later subjected their conclusions to independent peer review in demonstrating that alkaline mine drainage was a likely cause of biological degradation in Appalachian streams (Cormier et al., 2013a).

EPA considered potential confounding factors using an unusually large sample of sites across the region, including “habitat, organic enrichment, nutrients, deposited sediments, pH, selenium, temperature, lack of headwaters, catchment area, settling ponds, dissolved oxygen, and metals.” (p. 41). EPA found that only pH was a confounder and controlled it by removing sites with low pH. (p. 41). EPA concluded that “[t]he signal from conductivity was strong so that other potential confounders that were not strongly influential could be ignored with reasonable or greater confidence.” (p. 41). Once again, later independent reviews corroborated the interpretation that this observed relationship persisted even after confounding factors including habitat quality, deposited sediment, pH, selenium, catchment area, settling ponds and metals were considered and their effects addressed analytically (Cormier and Suter, 2013).

DRAFT

EPA's benchmark report also analyzed the relationship between the WVSCI biological impairment threshold and conductivity levels, and found that a WVSCI score of 64 (well below the current impairment threshold of 72) corresponds to streams with conductivity of about 300. (p. A-36). Since the WVSCI was developed independent of the benchmark, this is a separate method that validates the relationship between impairment and conductivity.

Post-Benchmark Studies. After the EPA benchmark was issued, a relevant corroborating study came from the same physiographic region but across state lines in Virginia. Here the authors used a selection of 22 sites where temperature and habitat conditions were of reference quality to test for the effect of total dissolved solids and conductance on macroinvertbrate indices (Timpano et al. 2011). The authors found that elevated conductance above 332 µS/cm associated with surface mining caused degradation for the VSCI and that levels below 465 uS/cm led to the occurrence of only 95% of reference genera. The results are important because they demonstrated that high conductivity levels alone are sufficient to generate degradation of biotic integrity in the absence of other stressors, and because those results are qualitatively very similar to those of US EPA (2011).

Also after the Benchmark, Palmer and Bernhardt published a peer-reviewed study in 2011 in the *Annals of the New York Academy of Sciences*. The report stated that surface mining in Central Appalachia has caused greatly increased sulfate concentrations and electrical conductivity in downstream waters, and that analysis of the West Virginia database of small streams "found that sulfate concentrations were highly correlated with conductivity, Ca, Cl, Fe, Mg, and Hardness—all of which contribute to heightened ionic stress in these impacted streams." (pp. 47-48). The report further found that this elevated conductivity leads to the loss of sensitive macroinvertebrate taxa, such as mayflies in Central Appalachian streams below coal mines. (p. 48)

Lindberg and Bernhardt published a peer-reviewed study in 2011 in the *Proceedings of the National Academy of Sciences*. The study found that all tributaries draining mountaintop-mining-impacted catchments in a portion of the Upper Mud River watershed in West Virginia were characterized by high conductivity and increased sulfate concentration. Sulfate concentration "was significantly positively correlated with constituents typically derived from rock and coal weathering (SO4, Ca, Mg, Li, Rb, and U) in the mainstem as well as the MTM-affected tributaries." (p. 2) The study "conclusively demonstrates that the observed increases in conductivity and Se concentration can be attributed directly to the area extent of surface coal mining occurring in the watershed." (p. 5) The study also stated that "the constituent weathering-derived salts that contribute to conductivity are not ameliorated nearly two decades after reclamation." *Id*.

Northington et al. (2011) published a paper in *Hydrobiologia* that examined the impacts of habitat restoration on streams draining alkaline mines in southeastern Virginia. They found no

DRAFT

impact of restoration on biotic impairment; in many cases, habitat scores were further degraded by restoration attempts and showed no impact on high levels of specific conductivity.

In 2012, Pond published a peer-reviewed paper in *Hydrobiologia* that showed species composition changed dramatically as a function of land use and that conductivity was an excellent indicator of how many individuals of certain types of macroinvertebrate taxa normally abundant in Appalachian streams would be found at a disturbed site. Pond compared types of land disturbance at 94 sites in Kentucky, including mining sites, and stated in the abstract that "[c]ore caddisfly genera were extirpated from most disturbed sites." He found that "no habitat factors were significantly correlated with relative abundance metrics," while "major ion concentrations (measured as specific conductance) were also highly correlated with Plecoptera and Trichoptera richness . . . but not abundance." *Id*. at 11-12. Average site tolerance value "was most strongly correlated with specific conductance." *Id*. He concluded that the predominant naturally occurring stonefly genera in eastern KY headwater streams "serve to indicate 'healthy' Appalachian streams" and his data "revealed high rates of extirpation of many genera and entire families from headwater streams affected by varying levels of mining and residential disturbance." *Id*. at 18.

In 2012, Pond et al. published a peer-reviewed paper in *Environmental Monitoring and Assessment*, the abstract of which "described the development, validation, and application of a geographically- and seasonally partitioned genus-level index of most probable stream status (GLIMPSS) for West Virginia wadeable streams." He found that GLIMPSS detected greater stream impacts to benthic invertebrates than did the WVSCI method because it used a more sensitive genus-level rather than a family-level analysis. The threshold for impairment as measured by GLIMPSS is a score of 55 for the Mountain Summer category and 53 for Mountain Spring Pond, Table 8, p. 1532. (The Mountains category has previously been used by the WVDEP to calculate GLIMPSS Scores for streams at issue in this case.) This paper also underscored the importance of direct interpretation of taxa lists and submetric indices when interpreting benthic responses to mountaintop mining.

In their 2012 peer-reviewed "How Many Mountains" study, Bernhardt and King found that streams receiving water from mining catchments had significantly higher conductivity than streams in unmined areas. They also found that, after screening out potential confounding factors, high conductivity was highly correlated with lower numbers of sensitive taxa and declining WVSCI scores. The study used a different statistical method than the method used in EPA's benchmark and reached the same conclusion that five percent of stream taxa are lost when conductivity reaches about 300 µS/cm. The study stated in its abstract:

> The extent of surface mining within catchments is highly correlated with the ionic strength and sulfate concentrations of receiving streams. Generalized additive models were used to estimate the amount of watershed mining, stream ionic

> strength, or sulfate concentrations beyond which biological impairment (based on state biocriteria) is likely. We find this threshold is reached once surface coal mines occupy >5.4% of their contributing watershed area, ionic strength exceeds 308 µS/cm−1, or sulfate concentrations exceed 50 mg/L−1. Significant losses of many intolerant macroinvertebrate taxa occur when as little as 2.2% of contributing catchments are mined.

The Bernhardt et al. (2012) paper also identified 50 taxa that consistently declined in response to conductivity associated with surface mining. The greatest cumulative taxa declines occurred at 283 uS/cm (95% CI 178-289) and 50 mg/L sulfate (95% CI 27-57). The taxa most sensitive to mining included a wide variety of mayfly, stonefly, caddisfly, and beetle larvae characteristic of Central Appalachian streams. Mayflies in particular appeared especially sensitive to ionic stress from alkaline mine drainage, though stoneflies and caddisflies are also vulnerable (see also Pond 2010, 2012). As expected, a few highly tolerant taxa increased in relative abundance along the mining gradient, primarily genera of highly tolerant midges (*Chironomidae*) and the tolerant caddisflies *Chimarra* and *Hydropsyche*.

As mentioned earlier, Cormier and Suter published six peer-reviewed studies in 2013 based on different sections of EPA's benchmark report in *Environmental Toxicology and Chemistry*, a high-quality scientific journal. In the first study entitled "A Method for Deriving Water-Quality Benchmarks Using Field Data," they described a method for using biological and water-quality parameters to develop a field-based benchmark to protect 95% of the genera from extirpation. The use of field data is helpful where lab-based data is not available, such as where susceptible species and sensitive life stages are difficult to maintain and test in the laboratory.

In the second study entitled "Derivation of a Benchmark for Freshwater Ionic Strength," they developed an aquatic life benchmark in West Virginia for specific conductance as a measure of ionic strength that is expected to prevent the local extirpation of 95% of species from neutral to alkaline waters containing a mixture of dissolved ions in which the mass of $SO2^{-4}$ + HCO^{-3} is greater than or equal to Cl-. Extirpation concentrations of specific conductance were estimated from the presence and absence of benthic invertebrate genera from 2,210 stream samples in West Virginia. The study concluded that the extirpation concentration is 300 µS/cm. One of the reasons for using field data rather lab data is that *Ephemeropterans* (mayflies), which are the most sensitive to the ionic mixture, are not available as cultured animals for toxicity tests.

In the third study entitled "A Method for Assessing Causation of Field Exposure-Response Relationships," Cormier and Suter developed a weight-of-evidence method to determine how an association in the field is causal. They identified six characteristics of causation: co-occurrence, preceding causation, interaction, alteration, sufficiency, and time order.

In the fourth study entitled "Assessing Causation of the Extirpation of Stream Macroinvertebrates by a Mixture of Ions," they applied that method to determine that the

DRAFT

relationship between conductivity and extirpation of benthic macroinvertebrates was causal. They stated in their abstract that "a mixture containing the ions Ca^{+}, Mg^{+}, HCO^{-3}, and SO^{-4}, as measured by conductivity, is a common cause of extirpation of aquatic macroinvertebrates in Appalachia where surface coal mining is prevalent."

In the fifth study entitled "A Method For Assessing The Potential For Confounding Applied To Ionic Strength In Central Appalachian Streams," they evaluated twelve potential confounders: habitat, organic enrichment, nutrients, deposited sediments, pH, selenium, temperature, lack of headwaters, catchment area, settling ponds, dissolved oxygen, and metals. They concluded that pH, temperature, habitat, and deposited sediments were not confounding factors.

In the sixth study entitled, "Relationship of Land Use and Elevated Ionic Strength in Appalachian Watersheds," they found that, based on a 10th quantile regression analysis, 300 µS/cm was exceeded when 3.3% or more of an area was covered by valley fills. They also confirmed that coal-mining activities are the primary source of high conductivity waters.

In 2014, Pond et al. published a study in *Environmental Management* that sampled fifteen headwater streams with valley fills in Central Appalachia that had been reclaimed from eleven to thirty-three years earlier. The study found that nearly 90% of these streams exhibited biological impairment, and that valley fill sites with higher WVSCI scores were located near undisturbed tributaries that could be the sources of sensitive taxa as drifting colonists. This result could explain why there are occasional passing stream condition index scores at sites when water chemistry and upstream land use would predict impairment. Pond stated in this article:

> Although these VFs were constructed pursuant to permits and regulatory programs that have as their stated goals that (1) mined land be reclaimed and restored to its original use or a use of higher value, and (2) mining does not cause or contribute to violations of water quality standards, we found sustained ecological damage in headwaters streams draining VFs long after reclamation was completed.

His three main conclusions were that "(1) temporal ecological impacts persist downstream of VFs, given 11-33 years post-reclamation; (2) many expected taxa were missing from VF streams (suggesting local extirpations) and the scraper feeding group was significantly reduced; and (3) water quality is most likely the primary barrier to recovery but proximity to clean sources (intervening tributaries) may contribute some sensitive taxa that increase the biological indices used to measure condition." *Id*. at 11. Elaborating on these three points, he further explained that conductivity was persistent and habitat was not a confounding factor for the observed stream impairment:

> our data indicated that highly elevated ionic concentrations may persist for over 30 years post-reclamation and that these chemical signatures result in damaged

DRAFT

> aquatic communities. Habitat can be a limiting factor, but by design, we removed significant habitat degradation factors by selecting sample reaches with relatively good habitat and intact riparian vegetation at reference and VF sites . . .
>
> after 11-33 years post-reclamation, bioassessment indices indicated persistent temporal effects; almost 90% of our streams draining old VFs scored below impairment thresholds using GLIMPSS and O/E [observed/expected predictive model]. . . .
>
> Overall, biological variation was strongly correlated with water chemistry and less with reach-scale habitat and landscape conditions. Since ion concentrations explained the greatest amount of biological impacts and were the most altered (compared to reference), this suggests that recovery is potentially hindered by ions, even in forested reaches long after reclamation. Causal analyses by Suter and Cormier (2013) provided evidence that ions (measured as specific conductance) negatively affected invertebrates despite other stressors present. . .
>
> Cormier et al. (2013b) and Suter and Cormier (2013) provided strong causal evidence that Appalachian macroinvertebrate extirpation is linked to increasing ions (as specific conductance), a finding supported by our study.

Id. at 12-13.

In 2014, USGS scientists Nathaniel Hitt and Douglas Chambers published a peer-reviewed paper looking at the effects of mountaintop mining on fish assemblages. Among other findings they noted that most obligate invertivores were extirpated at MTM sites, indicating that conductivity effects on macro-invertebrates resulted in impacts higher up the food chain, on fish. They also found that the effects of MTM were not related to physical-habitat conditions but were associated with water-quality variables, which may limit quality and availability of benthic macroinvertebrate prey.

More recently, several other authors have examined linkages between surface mining and biotic response using geographically extensive datasets. Voss et al. (2015) expanded on the work of Bernhardt et al. (2012) with their dataset and used Bayesian models of invertebrate abundance to generate a risk landscape of impairment at various levels of dissolved solids. They showed results that corroborated all earlier studies. Merriam et al. (2015a&b) used an independent sample of some 170 West Virginia streams that accounted for effects of surface mining, residential development, and underground mines on stream chemistry and biotic index scores. They showed that while all land uses produced increases in stream conductance, the chemical signature of each was relatively distinct, their impact on stream invertebrates was largely additive in streams with surface mining, and co-occurrence of residential development or

DRAFT

underground mines with surface mines could exacerbate the apparent effect of mining activity on conductivity and stream biota.

Cook et al. (2015) used a smaller set of nested sites from Virginia to assess the relative impact of conductivity and habitat degradation downstream of surface mines. They found a strong habitat degradation effect when habitat scores were quite low (i.e., "poor") and stream conductivity did not vary greatly. Bier et al. (2014) studied bacterial community responses to mountaintop mining below valley fills. In an article published in *The International Society for Microbial Ecology Journal*, they described marked changes in bacterial composition and, in some cases, functional capacity in streams with alkaline mine drainage.

In 2016, Ross et al. published a paper in *Environmental Science & Technology* that concluded the scale of surface mining impacts penetrates deep into bedrock and persists over longer time-scales than two-dimensional land use assessments would suggest. Rather than deforestation, the authors associated the effects of surface mining with impacts more closely associated with volcanic eruptions. Nippgen et al. (2017) published a paper in the same journal showing that the result of such deep impacts was an extension of perennial baseflow, shortening of elevated flow periods, and increases in flow yield from landscapes with valley fills when compared to those without valley fills. Such findings imply that in addition to elevated ionic concentrations, tributaries draining mined watersheds also contend with more constant levels of these concentrations through time, thus increasing the potential stress on stream biota. Such findings are stunning when combined with the temporal and spatial extent of surface mining mapped across Central Appalachia through time by Perciak et al. (2018) in the *Public Library of Science* (PLoS One).

In addition to their 2011 report, Timpano and his colleagues have conducted additional relevant work. In 2015, they published an article in the *Journal of the American Water Resources Association* that was a peer-reviewed recapitulation of their 2011 report. Boehme et al (2016) described seasonal patterns of impairment across a gradient of mining impacts in the peer-reviewed journal *Ecological Indicators*. They found altered benthic communities in medium to high specific conductivity streams marked by increases in conductivity-tolerant taxa and decreases in sensitive taxa. They found richness metrics to be more sensitive than aggregate indices or those based on relative abundance because some tolerant taxa occurred in most major groups. They proposed an alternate EPT formulation excluding relatively tolerant *Hydropsychids*, *Baetids*, and *Leuctrids*. They also showed that temporal biotic variability increased within medium conductivity streams as compared to reference or high conductivity streams. Timpano et al. (2018) published a follow-up paper, also in *Ecological Indicators,* in which they demonstrated seasonality in benthic response to alkaline mine drainage over a 4.5 year period. The authors identified groups of conductivity-sensitive and tolerant taxa and critical conductivity levels at which biotic metrics experienced significant changes in response to increasing conductivity.

Giam et al. (2018) found that streams affected by coal mining averaged one-third (32%) lower taxonomic richness and one-half (53%) lower total abundance than unmined streams, with these impacts occurring across all taxa investigated thus far (invertebrates, fish, and salamanders). Even after post-mining reclamation, biodiversity impact persisted. Giam identified the Elk River watershed in West Virginia as one of the watersheds of highest conservation concern because of reduced biodiversity from coal mining. The Seven Pines Mine discharges into tributaries of the Elk River.

Importantly, the analytical techniques employed by Bernhardt et al. (2012), Voss et al. (2015), Cook et al. (2015), and Merriam et al. (2015) differed markedly from those employed by US EPA (2011), Cormier et al. (2013c), Timpano et al. (2011; 2015; 2018), and Pond et al. (2008). Moreover, Pond collected samples from West Virginia and Kentucky; US EPA (2011), Bernhardt (2012), and Voss (2015) relied on the WVDEP sampling database; Merriam (2015) collected their own West Virginia data; whereas Cook (2015) and Timpano (2011, 2015, 2018) sampled in southwestern Virginia. Nonetheless, all groups arrived at qualitatively similar conclusions regarding the potential of degradation from stream conductance downstream of Appalachian surface mines. Many noted that when habitat is highly degraded, it too contributes to poor biotic condition. However, all agreed that conductivity is sufficient *by itself* to produce biotic impairment. Such agreement is rather rare in stream ecology, and supports interpretation of the general role of elevated conductivity as a causal agent in stream impairment.

Experimental Studies. Laboratory studies are also largely consistent with the tolerance thresholds of organisms derived from field observation. Kennedy et al. published a peer-reviewed paper in 2004 in *Environmental Monitoring and Assessment* which tested simulated coal mine discharge waters in Ohio. Seven-day lethality tests on *Isonychia bicolor*, a mayfly, found that lowest observed effect in mine effluent dominated by sulfates, bicarbonates, and sodium were 1582 µS/cm, 966 µS/cm, and 987 µS/cm in three tests. These values bracket the field-derived extirpation value of 1180 µS/cm.

In 2013, Kunz et al. published a peer-reviewed paper in *Environmental Toxicology and Chemistry.* Here, researchers exposed amphipods, mussels, and a species of mayfly to reconstituted mine water from 3 surface mines with an ionic composition characteristic of mountaintop-mining-impacted streams in Central Appalachia. Toxicity to the mayfly was determined to be between 800 and 1300 µS/cm, which is consistent with the field-derived extirpation value of 1092 µS/cm.

Clements & Kotalik (2016) were the first to be able to rear a healthy native stream assemblage complete with mayflies in a mesocosm experiment. They exposed stream assemblages to three types of effluent at various concentrations, and their results suggest that mining effluent similar to that in the Central Appalachian surface mines in southwestern Virginia is not immediately toxic to many invertebrates, but has its strongest impact on the survival of early instars while inducing changes in foraging behavior in older larvae. The authors concluded that the benchmark standards of the EPA (2011) were reasonable for protecting aquatic life.

Moving from mesocosms to natural experiments, Voss and Bernhardt (2017) published a study in *Limnology and Oceanography* that examined macroinvertebrate populations in an unimpacted stream as it conflued with mine-impacted tributaries. They found that not only were mining impacts associated with elevated levels of conductivity and sulfate as well as loss of sensitive taxa, the losses translated directly into depressed biomass throughout the year that were most apparent when pollutant concentrations rise with summer baseflow. They concluded that elevated ionic strength depresses insect production by preventing sensitive taxa from completing their life cycles in mining-impacted streams. Due to the domination of surface coal mining in Appalachia (recently documented by Perciak et al. 2018), altered production patterns are likely impacting regional food webs.

Summary of Scientific Research to Date. Together with the Benchmark, dozens of scientists in the field of ecology and ecological causation have reviewed the evidence establishing that conductivity in mine drainage is a cause of biological degradation in Appalachian streams. All of the science has passed peer review or the EPA's Scientific Advisory Board. The studies used a scientifically valid method of causal assessment. The primary data source used by EPA and Cormier et al. for evidence of confounding is West Virginia's watershed analysis database, which means that it is highly relevant to streams in southern West Virginia and neighboring states within the same region. The weight of evidence indicates that habitat, temperature, and sedimentation are not confounding factors in Central Appalachian mine sites generally or in this case specifically. There are no peer-reviewed studies with reliable study designs that contradict any of these findings.

The studies clearly show that levels of conductivity above ~300 uS/cm and elevated sulfate levels are common below Appalachian mine sites and lead to extirpation of invertebrate genera (EPA 2011; Cormier and Suter 2013; Cormier et al. 2013a; Timpano et al. 2015, 2018) and that the ions found coming out of the outlets at the Seven Pines Mine are consistent with those associated with coal mining pollution in this region (Pond et al. 2008; Palmer et al. 2010; Bernhardt and Palmer 2011; Lindberg et al. 2012;Pond et al. 2012; Pond et al. 2013; Pond et al. 2014; Timpano et al. 2015, 2018). The ionic mixture of calcium, magnesium, sulfate, and bicarbonate in circum-neutral mine water causes the loss of aquatic macroinvertebrates in Appalachian areas where surface coal mining is prevalent; it is the mixture of ions that causes the biological impairment (Cormier et al. 2013b; Comier and Suter 2013). These ions also lead to reductions in fish assemblages in the affected streams (Hitt et al. 2014).

Altogether, *nine* different scientific methods have been used in these different studies by different scientists to reach the same conclusion about the causal link between conductivity and downstream impairment. First, the Benchmark used a species sensitivity distribution to model the conductivity level at which different genera are extirpated, and determined that 5% of taxa are lost at 300 µS/cm (pp. 18-19). Second, the Benchmark modeled conductivity against WVSCI scores, and determined that 300 µS/cm corresponded to a failing WVSCI score of 64 (p. A-36). Third, the Benchmark used a logistic regression, and found that the probability of impairment, as measured by WVSCI, was 59% at 300 µS/cm and 72% at 500 µS/cm (p. A-36). Fourth, Timpano et al. (2011; 2015) used a controlled study design to isolate the effects of elevated conductivity from habitat and temperature. Fifth, King and Bernhardt in their paper used Generalized Additive Models (GAMs) and a different statistical method called Threshold Indicator Taxa Analysis (TITAN) to corroborate earlier results. Sixth, Voss used a Bayesian estimation procedure on the WVDEP data. Seventh, Timpano et al. (2018) used a multivariate approach (NMS) and GAMs fit to submetrics from independent samples in Virginia to reach qualitatively similar conclusions. Eighth, Clements & Kotalik (2016) used a mesocosm to demonstrate experimental effects of conductivity on aquatic insects. Finally, Voss and Bernhardt (2017) used an *in situ* natural experiment to demonstrate the impacts of high-conductivity mine drainage on mayfly production and survival.

Collectively, these papers show remarkably consistent results across researchers and analytical methods, when the enterprise of scientific publication and peer review is set up to reward dissent or critical interpretations. This pattern of consistent support and corroboration without any substantive contradiction from independent investigators constitutes extremely strong empirical evidence that ionic stress produced by surface coal mining in West Virginia is a general cause of biological impairment.

<u>Judicial Decisions</u>

In multiple opinions since June 2014, Judge Robert C. Chambers of the U.S. District Court for the Southern District of West Virginia (Huntington Division) has found as fact that high levels of ionic pollution, measured as conductivity, causes biological impairment in Appalachian streams. Further, Judge Chambers has recognized that conductivity is the result of mining operations such as the one proposed by the applicant. His findings have been upheld by the United States Court of Appeals for the Fourth Judicial Circuit.

In a June 2014 decision Judge Chambers wrote:

> The Court will now assess the evidence presented at trial to determine whether Plaintiffs have proven this aspect of their case by a preponderance of the evidence. First, it is important to note that the EPA has spoken to both general causation theories 1) through its October 2005 "Mountaintop Mining/Valley Fills in Appalachia Final Programmatic Environmental Impact Statement" ("EIS") and, most importantly, 2) through its March 2011 Benchmark, entitled "A Field–Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams." Pls.' Ex. 9; *see* Tr. at 61–62. In its EIS, the EPA identified two downstream impacts from mountaintop mining valley fills: 1) increases in conductivity and 2) decreases in the number of invertebrate taxa. *See* Tr. at 62.

In its nearly three-hundred page scientific Benchmark—after considering and then ruling out the potential confounding effects of habitat, organic enrichment, nutrients, deposited sediments, pH, selenium, temperature, lack of headwaters, catchment areas, settling ponds, dissolved oxygen, and metals—the EPA found that "salts, as measured by conductivity, are a common cause of impairment of aquatic macroinvertebrates" in Central Appalachian streams. EPA's Benchmark at A–1, B–1; *see also id.* at A–40 ("This causal assessment presents clear evidence that the deleterious effects to benthic invertebrates are *caused by, not just associated with,* the ionic strength[, i.e., conductivity,] of the water.... When [other potential] causes are absent or removed, a relationship between conductivity and ephemeropteran [, i.e. mayfly,] richness is still evident." (emphasis added)); *id.* at A–37 ("As conductivity increases, the occurrence and capture probability decreases for many genera in West Virginia ... at the conductivity levels predicted to cause effects. The loss of these genera is a severe and clear effect."). The Benchmark also found that "of the [nine] land uses ... analyzed, only mining especially associated with valley fills[, i.e., mountaintop mining with valley fills,] is a substantial source of the salts that are measured as conductivity." *Id.* at A–18.

The Benchmark ultimately concluded that the "chronic aquatic life benchmark value for conductivity" in West Virginia streams is 300 µS/cm. *Id.* at xv. To derive this recommended high-end threshold value, the EPA used the 5th percentile of a species sensitivity distribution, based on the standard methodology for deriving water-quality criteria, meaning that this 300 µS/cm benchmark value for conductivity is "expected to avoid the local extirpation [due to the salts measured as conductivity] of 95% of native species." *Id.* at xiv. In support of both the specific 300 µS/cm benchmark value and the general causal linkage between conductivity and impairment to aquatic macroinvertebrates, the Benchmark contains a graph which charts, for 163 genera, the level of salt exposure above which a genus is effectively absent from water bodies in a region, with conductivity readings on the x axis and proportion of genera extirpated on the y axis. *Id.* at xiv, 18 fig. 8. A fairly consistent line is formed as conductivity and extirpation both increase, illustrating the causal connection between conductivity and significant biological impairment which Plaintiffs seek to prove. *See id.* at 18 fig. 8. . . .

Second, two of the authors of the Benchmark, Dr. Susan Cormier and Dr. Glenn Suter, subsequently published four different peer-reviewed journal-article versions of several sections of the EPA's Benchmark—including the section regarding the causal link between conductivity and biological impairment and the section ruling out potential confounding factors—in the scientific journal *Environmental Toxicology and Chemistry. See* Tr. at 84. Plaintiffs' expert Dr. Palmer testified that this is a quality journal which focuses specifically on topics such as biological

DRAFT

response to pollutants. *Id.* at 84–86. Plaintiffs' expert Dr. King testified that "the list of the number of people who commented on [these journal articles] in the acknowledgments [section and] the peer reviews ... [is] impressive." *Id.* at 258. He also testified that, in his own professional opinion, he found the articles "rigorous and very defensible." *Id.*

Third, numerous other scientific articles published in peer-reviewed journals-both before and after publication of the Benchmark—lead to the same conclusions. In 2008, Dr. Gregory Pond-who would later be one of the contributors to the EPA's Benchmark—published a peer-reviewed scientific article in the *Journal of the North American Benthological Society,* based upon a field study he conducted which found that, as surface coal mining with valley fills—and its associated conductivity—increased, benthic macroinvertebrate taxa decreased. *See* Gregory J. Pond et al., *Downstream Effects of Mountaintop Coal Mining: Comparing Biological Conditions Using Family— and Genus–Level Macroinvertebrate Bioassessment Tools,* 27 J.N. Am. Benthological Soc'y 717 (2008), Pls.' Ex. 15; Tr. at 64–65; Pls.' Ex. 9. Dr. Palmer testified that the *Journal of the North American Benthological Society* is the highest impact freshwater journal in existence. Tr. at 65. . . .

Fourth, multiple different scientific methods were used at different times by different scientists to come to the same conclusions regarding the causal link between surface mining, conductivity, and biological impairment, which, Dr. Palmer testified, is the "strongest form of evidence" possible. Tr. at 83, 89–90, 248–52, 272, 274. For example, in its Benchmark, the EPA created a species sensitivity distribution—modeling the conductivity level at which each of 163 different genera are extirpated—which revealed that about five percent of taxa are lost at about 300 µS/cm. *See* EPA's Benchmark at 18–19. The Benchmark also used another method: modeling conductivity against WVSCI scores. *Id.* at A–35, –36. That modeled relationship revealed that the benchmark threshold of 300 µS/cm corresponded with a failing WVSCI score of 64. *Id.* at A–36. Using logistic regression, the probability of impairment—as measured by WVSCI—at 300 µS/cm was calculated to be 59%. *Id.* At 500 µS/cm, the probability of impairment was 72%. *Id.;* Tr. at 324. In the 2012 Bernhardt and King paper, two different methods were used to determine the biological impairment effects of conductivity: generalized additive regression models for three different biological response variables—including the number of intolerant taxa and WVSCI scores—and the Threshold Indicator Taxa Analysis ("TITAN") method, which Dr. King developed. *How Many Mountains* at C–D, P022–23; Tr. at 272, 274. Each of these different methods, conducted by different scientists at different times and subjected to the rigorous peer-review process required by scientific journals, resulted in the same conclusion: conductivity associated with surface mining

DRAFT

causes biological impairment, such that about five percent of taxa are lost at about 300 µS/cm. EPA's Benchmark at 18, A–36; *How Many Mountains* at F; Tr. at 274.

Fifth, the Court finds the expert testimony of Dr. Palmer and Dr. King to be very persuasive. . .

Finally, even though the WVDEP's Guidance purports to find that there is no causative effect between conductivity and low WVSCI scores, two portions of the Guidance seriously undermine this assertion and, ironically, support Plaintiffs' case. First, the Guidance includes a scatterplot graph of conductivity and associated WVSCI scores which reveals a clear reduction in WVSCI scores as conductivity increases; in fact, above 1500 µS/cm, only 2 scores out of approximately 100 fall above the passing WVSCI score threshold of 68 and the vast majority fall under 60.6. WVDEP's Guidance at 6. This strong association supports, rather than contradicts, a causal connection. Second, Figure 2 in the Guidance concludes that conductivity measurements that fall within the range of 1075–1532.9 µS/cm are “likely stressor[s]” and that measurements above 1533 µS/cm are “definite stressor[s].” *Id.* at 7. Almost all of the recent conductivity measurements at the sites at issue in this case fall within these two categories; many are firmly within the “definite stressor” category. Thus, the WVDEP's Guidance is additional evidence that high levels of conductivity cause biological impairment.

In the face of such overwhelming scientific evidence, this Court **FINDS** that Plaintiffs have proven, by a preponderance of the evidence, that, 1) controlling for other potential confounding factors, high conductivity in streams causes or at least materially contributes to a significant adverse impact to the chemical and biological components of aquatic ecosystems—proof of which can be shown through low WVSCI scores—and 2) surface mining causes—or at least materially contributes to—high conductivity in adjacent streams.[1]

Judge Chambers built upon his findings in a subsequent decision in a similar case issued in January 2015. There the Judge concluded:

In multiple ways, the chemical and the biological components of the aquatic ecosystems found in Stillhouse Branch have been significantly adversely affected by Defendant's discharges. The water chemistry of this stream has been dramatically altered, containing levels of ionic salts—measured as conductivity—which are scientifically proven to be seriously detrimental to aquatic life. The biological characteristics of the stream have also been significantly injured, in that species diversity—and, in some areas, overall aquatic life abundance—is profoundly reduced. Stillhouse Branch is unquestionably biologically impaired, in violation of West Virginia's narrative water quality

> standards, with current WVSCI scores falling well below the threshold score of 68.
>
> Losing diversity in aquatic life, as sensitive species are extirpated and only pollution-tolerant species survive, is akin to the canary in a coal mine. This West Virginia stream, like the reference streams used to formulate WVSCI, was once a thriving aquatic ecosystem. As key ingredients to West Virginia's once abundant clean water, the upper reaches of West Virginia's complex network of flowing streams provide critical attributes—"functions," in ecological science—that support the downstream water quality relied upon by West Virginians for drinking water, fishing and recreation, and important economic uses. Protecting these uses is the overriding purpose of West Virginia's water quality standards and the goal of the state's permit requirements.[2]

Judge Chamber's decision in that case was ultimately upheld in full, by the United States Court of Appeals for the Fourth Judicial Circuit. After examining his review of the evidence and the applicable law, that court found:

> The court noted that peer-reviewed scientific articles first recognized the relationship of mining, conductivity, and decreased Index scores in 2008, a year before issuance of [the permittee's] renewal permit. See [Chambers Decision] (citing Pond et al., supra n.1). Other articles strengthened these findings. *Id.* (citing, among others, M.A. Palmer et al., Mountaintop Mining Consequences, 327 Sci. 148 (2010) (finding that as conductivity increased, Index scores decreased)). In rebuttal, [the permittee] offered an expert whom the district court found unqualified—an assessment [the permittee] does not challenge on appeal. . . . In sum, [the permitte's] arguments as to why the district court erred in finding that [it] violated its permit, like [the permitee's] arguments as to the permit's reach, uniformly fail.[3]

Judge Chambers issued another decision in August 2015, finding:

> On the basis of this outstanding collection of peer-reviewed studies, the Court finds that the link between surface mining and biological impairment of downstream waters has been sufficiently—if not definitively—established in the scientific literature. "There's field data. There's lab data. There's observational data. There's field experimental data. There's toxicity testing." Tr. 2 at 141, ECF No. 100. Through myriad lines of evidence, researchers have reached the same general causation conclusion, without a single peer-reviewed publication reporting contrary findings. In Dr. Palmer's expert opinion, there is no remaining doubt on the question of general causation, leaving only surprise that researchers are continuing to study the question. *Id.* at 141 ("I would say there's no doubt.

DRAFT

What surprised me is that the studies continue to go on.... because it's been so well-established.”).[4]

Subsequent cases have continued to result in findings that ionic toxicity measured as conductivity *causes* biological impairment in streams beneath surface mines.[5]

References

Bernhardt, E. S., B. D. Lutz, R. S. King, A. M. Helton, C. A. Carter, J. P. Fay, D. Campagna, J. Amos. 2012. How many mountains can we mine? Assessing the regional degradation of Central Appalachian rivers by surface coalmining. *Environmental Science & Technology* 46: 8115–8122.

Bier, R. L., Voss, K. A., & Bernhardt, E. S. (2014). Bacterial community responses to a gradient of alkaline mountaintop mine drainage in Central Appalachian streams. *The ISME journal*, *9*(6), 1378.

Clements, W. H., & Kotalik, C. (2016). Effects of major ions on natural benthic communities: an experimental assessment of the US Environmental Protection Agency aquatic life benchmark for conductivity. *Freshwater Science*, *35*(1), 126-138.

Cook, et al, (2015) Habitat and water quality as drivers of ecological system health in Central Appalachia. *Ecological Engineering* 84:180-189.

Cormier S. and G.W. Suter. 2013. A method for assessing causation of field exposure- response relationships. *Environmental Toxicology and Chemistry* 32:272–276.

Cormier SM, Suter, GW, Pond GJ, Zheng L. 2013a. Assessing causation of the extirpation of stream macroinvertebrates by a mixture of ions. *Environmental Toxicology and Chemistry* 32:277–287.

DRAFT

Cormier S., S. P. Wilkes, and L. Zheng. 2013b. Relationship of land use and elevated ionic strength in Appalachian watersheds. *Environmental Toxicology and Chemistry* 32:296-303.

Cormier SM, Suter GW, and Zheng L. 2013c. Derivation of a benchmark for freshwater ionic strength. *Environmental Toxicology and Chemistry* 32:263–271.

Cormier, S. M., Suter, I. I., & Glenn, W. (2013). A method for deriving water-quality benchmarks using field data. *Environmental Toxicology and Chemistry* 32(2), 255-262.

Gerritsen, J., Zheng, L., Burton, J., Boschen, C., Wilkes, S., Ludwig, J., & Cormier, S. M. (2010). Inferring causes of biological impairment in the Clear Fork Watershed, West Virginia. EPA 600/R-08/146. US Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, Cincinnati, OH, USA.

Giam, X, Olden, J.D., and Simberloff, D. (2018). Impact of coal mining on stream biodiversity in the U.S. and its regulatory implications. *Nature Sustainability* 1:176-183.

Griffith, M. B., Norton, S. B., Alexander, L. C., Pollard, A. I., & LeDuc, S. D. (2012). The effects of mountaintop mines and valley fills on the physicochemical quality of stream ecosystems in the central Appalachians: A review. *Science of the Total Environment* 417, 1

Hitt, N. P., & Chambers, D. B. (2014). Temporal changes in taxonomic and functional diversity of fish assemblages downstream from mountaintop mining. *Freshwater Science* 33(3), 915-926.

Hobbs, R. J., Arico, S., Aronson, J., Baron, J. S., Bridgewater, P., Cramer, V. A., ... & Zobel, M. (2006). Novel ecosystems: theoretical and management aspects of the new ecological world order. *Global ecology and biogeography*, 15(1), 1-7.

DRAFT

Kalisch, M., Mächler, M., Colombo, D., Maathuis, M. H., & Bühlmann, P. (2012). Causal inference using graphical models with the R package pcalg. *Journal of Statistical Software*, *47*(11), 1-26.

Kennedy, A. J., Cherry, D. S., & Currie, R. J. (2004). Evaluation of ecologically relevant bioassays for a lotic system impacted by a coal-mine effluent, using Isonychia. *Environmental Monitoring and Assessment* 95(1-3), 37-55.

King, RS and ME Baker. 2014. Use, misuse, and limitations of Threshold Indicator Taxa Analysis (TITAN) for estimating ecological community thresholds. In: G. Guntenspergen (editor), Application of Threshold Concepts in Natural Resource Decision Making, Springer, New York.

King, R. S., M. E. Baker, P. F. Kazyak, and D. E. Weller. 2011. How novel is too novel? Stream community thresholds at exceptionally low levels of catchment urbanization. *Ecological Applications* 21:1659-1678

Kunz, J. L., Conley, J. M., Buchwalter, D. B., Norberg, King, T. J., Kemble, N. E., Wang, N., & Ingersoll, C. G. (2013). Use of reconstituted waters to evaluate effects of elevated major ions associated with mountaintop coal mining on freshwater invertebrates. *Environmental Toxicology and Chemistry* 32(12): 2826-2835.

Lindberg, T. T., Bernhardt, E. S., Bier, R., Helton, A. M., Merola, R. B., Vengosh, A., & Di Giulio, R. T. (2011). Cumulative impacts of mountaintop mining on an Appalachian watershed. *Proceedings of the National Academy of Sciences* 108(52): 20929-20934.

Merriam, E.R., JT Petty, GT Merovich, JB Fulton, and MP Strager. 2011. Additive effects of mining and residential development on stream conditions in a central Appalachian watershed. *Journal of the North American Benthological Society*, 30:399-418.

DRAFT

Merriam, E. R., Petty, J. T., Strager, M. P., Maxwell, A. E., & Ziemkiewicz, P. F. (2015). Landscape-based cumulative effects models for predicting stream response to mountaintop mining in multistressor Appalachian watersheds. *Freshwater Science*, *34*(3), 1006-1019.

Miller, Andrew & Zégre, Nicolas, (2016) Landscape-scale disturbance: Insights into the complexity of catchment hydrology in the mountaintop removal mining region of the eastern United States. *Land* 5(3):22.

Nippgen, F., Ross, M.R.V., Bernhardt, E.S., and McGlynn, B.L. (2017). Creating a more perennial problem? Mountaintop removal coal mining enhances and sustains saline baseflows of Appalachian watersheds. *Environmental Science & Technology* 51(15): 8324–8334.

Northington, R. M., Benfield, E. F., Schoenholtz, S. H., Timpano, A. J., Webster, J. R., & Zipper, C. (2011). An assessment of structural attributes and ecosystem function in restored Virginia coalfield streams. *Hydrobiologia*, *671*(1), 51-63

Pearl, J. (1988). Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann.

Pearl, J. (2000). Causality: models, reasoning and inference (Vol. 29). Cambridge: MIT press.

Perciak, A, Wasson MF, Ross MRV, et al. (2018) Mapping the yearly extent of surface coal mining in Central Appalachia using Landsat and Google Earth Engine. *PLOS ONE* 13(7): e0197758. https://doi.org/10.1371/journal.pone.0197758

Petty, JT, J.B. Fulton, M. P. Strager, G. T. Merovich, J.M. Stiles, and PF Ziemkiewicz. 2010. Landscape indicators and thresholds of stream ecological impairment in an intensively mined Appalachian watershed. *Journal of the North American Benthological Society* 29:1292–1309.

Pond, G. J. M.E. Passmore, F.A. Borsuk, L. Reynolds, and C.J. Rose. 2008. Downstream effects of mountaintop coal mining: comparing biological conditions using family- and genus- level macroinvertebrate bioassessment tools. *Journal of the North American Benthological Society* 27:717-737.

DRAFT

Pond, G. 2010. Patterns of Ephemeroptera taxa loss in Appalachian 3 headwater streams (Kentucky, USA). *Hydrobiologia* 641:185–201.

Pond, G. J. 2012. Biodiversity loss in Appalachian headwater streams (Kentucky, USA): Plecoptera and Trichoptera communities. *Hydrobiologia* 679(1):97-117.

Pond, G.J., J.E. Bailey, B.M. Lowman, and M.H. Whitman. 2013. Calibration and validation of a regionally and seasonally stratified macroinvertebrate index for West Virginia wadeable streams. *Environ. Monitoring Assess.* 185:1515–1540.

Pond, G. J., Passmore, M. E., Pointon, N. D., Felbinger, J. K., Walker, C. A., Krock, K. J., ... & Nash, W. L. 2014. Long-Term Impacts on Macroinvertebrates Downstream of Reclaimed Mountaintop Mining Valley Fills in Central Appalachia. *Environmental Management* 1-15.

Ross, M. R. V., Nippgen, F., Hassett, B. A., McGlynn, B. L., & Bernhardt, E. S. (2018). Pyrite oxidation drives exceptionally high weathering rates and geologic CO_2 release in mountaintop-mined landscapes. *Global Biogeochemical Cycles*, 32. https://doi.org/10.1029/2017GB005798

Ross, M. R., McGlynn, B. L., & Bernhardt, E. S. 2016. Deep Impact: Effects of Mountaintop Mining on Surface Topography, Bedrock Structure, and Downstream Waters. *Environmental Science & Technology* 50(4):2064-2074.

Shipley, B. (2002). Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference. Cambridge University Press.

Stow, C. A., & Borsuk, M. E. 2003. Enhancing causal assessment of estuarine fishkills using graphical models. *Ecosystems* 6(1):0011-0019.

Suter, G.W. and S.M. Cormier. 2013. A method for assessing the potential for confounding applied to ionic strength in central Appalachian streams. *Environmental Toxicology & Chemistry* 32:288-295.

DRAFT

Tetratech. 2000. A stream condition index for West Virginia wadeable streams. http://www.dep.wv.gov/WWE/watershed/bio_fish/Documents/WVSCI.pdf

Timpano, A. J., Schoenholtz, S., Zipper, C., & Soucek, D. (2011). Levels of dissolved solids associated with aquatic life effects in headwater streams of Virginia's Central Appalachian coalfield region (Doctoral dissertation, University Libraries, Virginia Polytechnic Institute and State University).

Timpano, A. J., Zipper, C. E., Soucek, D. J., & Schoenholtz, S. H. (2018). Seasonal pattern of anthropogenic salinization in temperate forested headwater streams. *Water research*, *133*, 8-18.

US EPA (Environmental Protection Agency). 2000. Stressor identification guidance document. US EPA (Environmental Protection Agency). 2011. A Field-based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams. Office of Research and Development, National Center for Environmental Assessment, Washington, DC. EPA/600/R-10/023F.

Voss, K. A., King, R. S., & Bernhardt, E. S. 2015. From a line in the sand to a landscape of decisions: a hierarchical diversity decision framework for estimating and communicating biodiversity loss along anthropogenic gradients. *Methods in Ecology and Evolution* 37(1):130-137.

Voss, K.V. and **E.S. Bernhardt.** 2017. Effects of mountaintop removal coal mining on the diversity and secondary productivity of Appalachian rivers. *Limnology and Oceanography 62*: 1754-1770.

Zuur, A. F., Ieno, E. N., & Elphick, C. S. 2010. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* 1(1):3-14.

DRAFT

[1] *Ohio Valley Environmental Coalition v. Elk Run Coal Co.*, 24 F.Supp.3d 532, 556–63 (S.D. W. Va. 2014) (internal footnotes omitted).

[2] *Ohio Valley Environmental Coalition v. Fola Coal Co.*, 82 F.Supp.3d 673, 698–99 (S.D. W.Va. 2015).

[3] *Ohio Valley Environmental Coalition v. Fola Coal Company, LLC,* 845 F.3d 133 (4th Cir. 2017).

[4] *Ohio Valley Environmental Coalition v. Fola Coal Co.*, 120 F.Supp.3d 509, 537 (S.D. W.Va. 2015).

[5] *Ohio Valley Environmental Coalition vs Fola Coal Co.*, 274 F.Supp.3d 378 (S.D. W.Va. 2017); *OVEC v. Lexington Coal Co.,* 2021 WL 1093631 (S.D. W.Va. March 22, 2021).

DRAFT