Published in final edited form as:

Sci Total Environ. 2018 July 01; 628-629: 1637–1649. doi:10.1016/j.scitotenv.2018.02.017.

Assessing background levels of specific conductivity using weight of evidence

Susan M. Cormier[a,*], **Lei Zheng**[b,1], **Glenn W. Suter II**[a], **Colleen M. Flaherty**[c]

[a]U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, Cincinnati, OH, USA

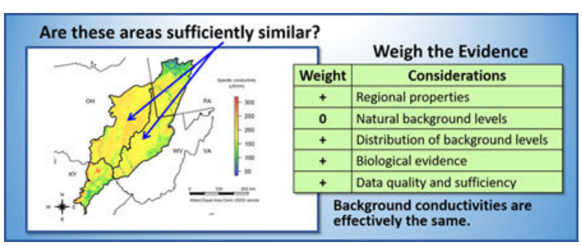
[b]Tetra Tech, 10711 Red Run Blvd., Suite 105, Owings Mills, MD 21117, USA

[c]U.S. Environmental Protection Agency, Office of Water, Office of Science and Technology, Health and Ecological Criteria Division, Washington, DC 20460, USA

Abstract

There are many ways to estimate background levels, and many types of evidence may contribute to determining whether a water, air, or soil is at background. As a result, it is important to define background in each case and to weigh the available evidence to determine the best estimate of background. A weight-of-evidence approach is demonstrated that assesses whether the background SC is sufficiently similar in streams of Ecoregion 70 in West Virginia and Ohio. During planning, five relevant considerations were identified to assess background SC: physical properties, measured SC, spatial distribution of low SC sites, biological properties, and data relevance and reliability. For each consideration, diverse types of evidence were generated, evaluated, and synthesized using weight of evidence. In the example, evidence was weighed for the hypothesis that background SC is similar in two areas in Ecoregion 70, the Western Allegheny Plateau in the eastern United States. Where, as in this case, background is not well characterized by measurements, because data sets are small or sampling designs or anthropogenic inputs may influence estimates of background, it is suggested that information about regional properties, related to and affected by SC, may be used to determine whether SC in the less characterized area is sufficiently similar to a well characterized area.

Graphical Abstract



*Corresponding author: cormier.susan@epa.gov (S.M. Cormier).

1Present address: Noblis, 2275-J rolling Run Dr., Woodlawn, MD 21244, USA.

Keywords

Background assessment; Salt; Geophysical Stream; Water quality criterion; Criterion applicability

1. Introduction

Environmental assessments require the estimation of background levels of a chemical or other environmental variable for several purposes. Background is defined as chemical concentrations or physical conditions that distinguish areas with anthropogenic disturbance from undisturbed areas. Background levels are specified to distinguish polluted sites from diffuse regional contamination or natural levels, to determine which areas should be remediated, and to determine who, if anyone, is responsible for observed levels (U.S. Environmental Protection Agency [U.S. EPA], 2002). In addition, they may be used in the derivation of criteria or benchmark values, which are concentrations that divide acceptable from unacceptable levels. In the European Union and some other jurisdictions, protective benchmarks are based on background as well as risk levels (European Union, 2000, 2006). Background may set lower limits on benchmark values, thereby restricting the magnitude of precautionary safety factors or assumptions. Because organisms are adapted to background levels, their tolerance of contamination by natural elements such as salts is related to the increment over background (U.S. EPA, 2016a; Cormier et al., 2018a). Because background implies these restraints on what concentrations may be considered contamination, defining areas with a common background level can determine the boundaries of applicability of a potential benchmark or remedial activity.

There are no generally accepted methods for defining or estimating background, but there are numerous options, which are associated with different concepts of background or different types of data (Reimann and Garrett, 2005; Reimann et al., 2005; Galuszka, 2007; Mast et al., 2007). Some of the options are:

1. Low levels for a chemical may represent background. For example, the U.S. EPA commonly has used the 25th centile of regional levels as the background level (EPA, 2000).
2. Anything that is not an anomaly (e.g. not an ore body, salt spring, or a stream reach receiving an effluent) may be background. This results in background limits such as, the mean plus two standard deviations (Reimann and Garrett, 2005).
3. When concentrations at background and contaminated sites have distinct distributions, background has been defined by inflections, break-points, or deviations from normality (Molinari et al., 2014). For example, the contaminant lead was distinguished from the natural concentration distribution in soil by deviation from linearity in a probability plot (Zhao et al., 2006).
4. Background may be what is found in uncontaminated or undisturbed sites or materials such as reference streams, deep sediments, ice cores, predevelopment analyses, or museum specimens. For example, Hinsby et al. (2008) used the 90th

or 97.5th centile of uncontaminated groundwater concentrations as background, depending on the amount and quality of data. The U.S. EPA commonly uses the 75th centile of reference site surface water concentrations as the limit of background, because some surface water reference sites are the best available (EPA, 2000).

5. Background may be the concentration in the source materials. For example, background for Rhine River sediments was assumed to be the concentration of elements in upper watershed soils (Van de Meent et al., 1990).

6. Background may be estimated by hydro-bio-geo-chemical models (Runnells et al., 1992; Mast et al., 2007). For example, background nutrient levels are empirically modeled from runoff and other watershed characteristics (Smith et al., 2003).

7. Background may be distinguished by differences in isotopic composition of natural and unnatural materials or differences in the "fingerprints" of chemical mixtures (the relative concentrations of constituents) from different sources. For example, polyaromatic hydrocarbon mixtures may differ among natural oil seeps, bunker fuels, and oil spills (Page et al., 1996). Similarly, stable isotope ratios have been used to determine natural and anthropogenic sources of lead (Sucharova et al., 2014; Luo et al., 2015).

8. If background may be defined by the absence of input from identified sources, it may be determined from the distribution of concentrations relative to the sources. For example, background elemental concentrations for soils receiving coal fly ash were defined by the asymptotic concentrations on sampling gradients away from a power plant (Gough and Crock, 1997). Alternatively, source strength, dilution, and losses may be modeled to determine what proportion of downstream concentrations is not background (Helgen and Moore, 1995).

9. Finally, background may be the level that best displays the characteristics of background given the body of evidence. After reviewing and dismissing individual methods to define background, Reimann et al. (2005) recommended graphical inspection of various maps and plots (particularly box plots) and application of integrative judgment. The USGS similarly applies multiple analyses to background derivation (Mast et al., 2007). Those approaches are limited to analyses of concentration data. This paper applies a formal weight-of-evidence (WoE) approach that includes other types of evidence as well as analyses of concentration data.

In this paper, background SC is defined as the range of SC naturally occurring in waters that have not been substantially influenced by human activity. A background value is the upper limit of the background range. Defining background and the boundaries within which background may be assumed to be uniform is complicated by three sources of variability: natural variance in geology, hydrology, and biology; variance in human input; and variance in sampling and analysis procedures. Because of the complexity of spatial and temporal variability in chemical concentrations, concentration data alone are often insufficient to estimate background or define areas of background condition (U.S. EPA, 2011a).

Because true natural background conditions may be rare for SC, natural background was approximated using minimally affected sites. Minimally affected conditions are the physical, chemical, and biological habitat found in the absence of significant human disturbance (Stoddard et al., 2006). In some cases, disturbance is so pervasive that an anthropogenic background is estimated from least disturbed conditions. These are the best available physical, chemical, and biological habitat conditions given the present degree of disturbance of the landscape or habitat type (Stoddard et al., 2006). In this paper, conditions did not necessitate the use of anthropogenic background.

This case study applies WoE to a condition assessment that determines whether background specific conductivity (SC) in one well-studied area is applicable to an adjoining area so that an aquatic life criterion may be applied to both areas. Criteria derived by a field-based method have been shown to be affected by background (U.S. EPA, 2016a; Cormier et al., 2018a, b). In such situations, several different types of evidence can be used to compare SC measurements in the two areas as well as other information concerning properties of the areas.

Natural background SC is likely to be more similar within an ecoregion than between ecoregions, because stream SC originates by climatic, geologic, and physiographic processes, which are also important determinates of ecoregions (Omernik, 1987). Portions of the same ecoregion in different political jurisdictions are expected to have similar characteristics with respect to the primary factors that control background SC (Hem, 1985; Olson and Hawkins, 2012; Griffith, 2014). The primary factors are underlying geology, physiography, and climate; secondary factors include soils and vegetative cover (Hem, 1985; Olson and Hawkins, 2012; Griffith, 2014). However, ecoregional boundaries have not been expressly developed as areas of similar water characteristics. Further, political boundaries may be associated with differences in sampling and analysis methods and timing that cause artificial differences in apparent background.

There may be situations where it is not appropriate to apply ecoregional background to an area or a particular stream reach. For example, naturally lower or higher concentrations of ions may occur due to sub-ecoregional differences such as cross boundary influences, glacial deposits, salt springs, highly soluble rock, or other local natural sources. Furthermore, the variability within an ecoregion may be too large for the intended application.

This paper describes how to weigh the evidence for background SC. The particular case determines whether the condition in a poorly characterized area is sufficiently similar to a well characterized area to treat both areas as having the same background.

2. Methods

2.1. Case study location

The case study is located in a well-characterized area in the southeastern portion of Ecoregion 70 within the state of West Virginia (original area) and the less characterized area is the northwest portion of Ecoregion 70 within the state of Ohio (new area) (Fig. 1). Initially, the new area in Ohio appeared to have a higher background based on a small data set compared to the well-characterized original area in West Virginia. If the two areas have different background SC, then it would not be appropriate to use a regional background encompassing both areas.

2.2. Condition assessment method

This assessment followed three basic steps that are inherent to any type of assessment: planning, analysis, and synthesis (Cormier and Suter, 2008). Planning included identification of assessment questions, relevant evidence, and properties of evidence that have implications for inferring background. During analysis, diverse types of evidence were generated for these considerations and weighted. Finally, the body of evidence was evaluated and synthesized using WoE (U.S. EPA, 2016b, Suter et al., 2017).

2.2.1. Planning method—Five relevant considerations were used to assess the similarity of background SC between two areas.

A. *Regional Properties*: Is the background SC expected to be similar in the new area and the original area based on knowledge of the regional physical, chemical, biological, and climatic properties that determine background SC?

B. *Empirically Estimated Background*: Is the statistical distribution of SC in the new area similar to the original area, particularly at the low end, which is indicative of background?

C. *Distribution of Low SC Sites*: Is the apparent background SC (based on available measurements) spatially dispersed across the new area?

D. *Biological Indicators of Background*: Are there past or present records for the new area of biota that inhabit only areas with a background similar to background in the original area and are therefore indicative of a similar background SC in both areas?

E. *Relevance and Reliability of Evidence*: Do the available data and other information for the new area provide sufficiently relevant and reliable evidence for assessing the background SC regime?

2.2.2. Analysis method—During the analysis phase, different types of evidence were generated and related to the five considerations for characterizing background SC and for assessing departure from the natural background SC regime. The types of evidence and their implications are listed in Table 1. They are appropriate to this case but may not be appropriate to other WoE assessments such as estimating background levels or mapping background patterns.

After the evidence was assembled, screened, and sorted by consideration, the logical implications and strength of the evidence were weighted. Considerations A–D are weighted with respect to the logical implications of the evidence (does it support similarity of the areas or not?) and the strength of the evidence. Consideration E addresses the relevance and reliability of the body of evidence. Evidence scored as + indicates that it supports similarity, − indicates that it weakens similarity, and 0 indicates that the evidence is absent or ambiguous. Multiple symbols indicate greater weight of the evidence. The information and weights of the evidence for each consideration were weighed and summarized in a WoE table.

Note that when applying this or any other WoE method, some evidence will be weak and other evidence will be strong. Evidence that fails acceptance criteria should have been screened out when the evidence was assembled. All remaining evidence is presented, even weak evidence, so that the full body of evidence can be given proper weight in determining the best-supported conclusion.

2.2.3. Synthesis method—Synthesis included evaluating and summarizing the scored evidence for each consideration and the body of evidence as a whole. We kept in mind the assessment questions, the potential causes for disparities or similarities, and the relevance for the proposed application of the assessment. To best explain the findings and implications of the evidence, the synthesis and interpretation of the body of the assessment relied upon accepted scientific principles, logic, and the professional experience of the researchers and reviewers. In this case, the synthesis must determine whether the weight of evidence supports or weakens the hypothesis that background is sufficiently similar in the two areas. It must then determine whether the body of evidence is sufficient or more information is needed.

3. Results

3.1. Example case study: Planning

The example assessment was initiated to determine whether background SC of streams in Ecoregion 70 in Ohio (new area) was sufficiently similar to the background in Ecoregion 70 in West Virginia (original area) to use a single regional background level to develop example water quality criteria or for other assessments that require estimates of background. The large number of high-quality SC values from the original area, including well defined reference sites, made its background level reliable. However, the minimum value from a very small probability survey ($N = 11$) of the portion of Ecoregion 70 in Ohio was 232 μS/cm, which is greater than the upper confidence limit for background (210 μS/cm) for the original area. To determine whether the estimates of background SC in the two portions of Ecoregion 70 were really distinct, the WoE method described in the method section was used.

3.1.1. Study area Western Allegheny Plateau (Ecoregion 70)—The Western Allegheny Plateau (Ecoregion 70) is a hilly, wooded plateau straddling the Ohio River (Omernik, 1987). Roughly centered at the cities of Parkersburg, WV and Marietta, OH, this diamond-shaped ecoregion covers southeastern Ohio and northwestern West Virginia and extends northward into southwestern Pennsylvania and southward into Kentucky (Fig. 1).

EPA Author Manuscript

EPA Author Manuscript

Bedded sedimentary sandstone, shale, siltstone, and limestone underlie this mostly unglaciated ecoregion. The area is mostly forested with smaller urbanized areas, pastures, farms, and surface and underground coal mines (Woods et al., 1996; Woods et al., 1999).

3.1.2. Study data sets—Multiple data sets were used in this case to assess similarity of background between the areas. To estimate background SC, the original area data set for Ecoregion 70 in West Virginia and the EPA-survey data set for all of Ecoregion 70 were used. Those data sets are described in U.S. EPA (2016a). An additional data set from the Ohio Environmental Protection Agency (Ohio EPA) that was used to estimate SC in the new area, is described here.

The Ohio EPA provided two data sets that were combined into a single new area data set. One data set (1999–2000, N=18) sampled small streams to develop a stream index for headwater streams. The second larger data set included wadeable sites sampled between 1999 and 2013 in Ohio (Fig. 2). This larger data set was designed to monitor conditions above and below outfalls of streams generally >50 km^2 catchment and water depth at least 10 cm at summer low flow to capture higher concentrations and stressful conditions.

Data filters were applied to the raw Ohio data set. First, samples from lakes, wastewater and industrial outfalls, and the Ohio River were excluded. Sites with pH <6 and one chloride-dominated site (where $[HCO_3^-] + [SO_4^{2-}] \leq [Cl^-]$) were also excluded. If repeat samples were taken at a location on the same date, only the minimum SC value was retained in the data set as most likely to reflect background. This new area data set contained 4452 stream samples from 809 unique stations (Fig. 2a). Repeat samples (81.8%) within a year at a unique station (1–15 times per annum) were used to estimate annual SC for a site (Table 2). Catchment areas ranged from 0.26 km^2 to 20,850 km^2 with 50% of sites sampled in drainages <32.1 km^2 and 8.5% >1500 km^2 (Table 3). Extensive documentation of field and laboratory methods is available from the Ohio EPA (http://www.epa.state.oh.us/dsw/bioassess/BioCriteriaProtAqLife.aspx). Unlike the EPA survey data set (probability survey designed) and the West Virginia data set for the original area (probability sampling plus some purposive sampling), the available Ohio data sets are based on Ohio EPA's geometric site selection process that stratifies a watershed based on a sequential, systematic halving of drainage area. Therefore, the data sets cannot be directly compared. They can, however, be used to determine whether low-end SC levels in Ohio are in the same SC range as similar sites in the rest of the ecoregion.

Analyses were performed using the new area data set because (1) it spanned Ecoregion 70 in Ohio (Fig. 2a), (2) the sampling window included samples throughout the year (Table 2), (3) it included headwater and wadeable streams making it more like the original area data set for West Virginia Ecoregion 70, and (4) multiple samples in a single year at a site enable some analyses of seasonality (Table 2).

The new area data set was used to assess background, annual average, and maximum SC of sites with $HCO_3^- + SO_4^{2-}$ concentrations on a mass basis greater than Cl^- (Table 4). Only 1 out of 601 sites in the new area data set was dominated by chloride and it had high SC (>1000 μS/cm). It was removed from the data set.

3.2. Example case study: Analysis

Data were analyzed, as far as possible, to generate each type of evidence (Table 5). R, Version 2.12.1 was used for all statistical analyses (R Development Core Team, 2011).

3.2.1. Evidence for consideration A: Regional properties—It is expected that areas within the same ecoregions in different states will have similar characteristics with respect to the properties that control background SC, but a natural gradient could lead to differences. The primary factors that affect natural SC are underlying geology, physiography, and climate. Secondary factors include soils and vegetative cover (Hem, 1985; Olson and Hawkins, 2012; Griffith, 2014; Anning and Flynn, 2014). Because ecoregions were delineated based on similar considerations (Omernik, 1987), the SC regime and ionic composition of dissolved salts in streams within an ecoregion tend to be similar throughout. Any degree of difference is affected by resolution of the geographical delineation and the homogeneity of the region. Level III ecoregions were judged to be a practical and reasonable level of aggregation for this approach.

3.2.1.1. Climate.: The climate of Ecoregion 70 in Ohio and in the other states is very similar. No large mountains or escarpments create elevational transitions or rain shadows between states and no large lakes create lake effect precipitation. The lake effect from Lake Erie does not extend into Ecoregion 70.

3.2.1.2. Geology and hydrology.: Bedded sandstone, shale, siltstone, limestone, and coal extend across both areas. The bituminous coal has been mined in both areas (Woods et al., 1996). The bedrock formations in Ecoregion 70 in West Virginia are nearly a mirror image of those in Ohio (Schruben et al., 1997; ODGS, 2006; WVGES, 2011). The bedrock along the Ohio River (the border between the areas) is transitional Pennsylvanian-Permian and the adjoining areas on both sides are Pennsylvanian.

Groundwater associated with the weathered zone of rock has low dissolved solids, because readily soluble products have been removed by chemical weathering (Brady, 1998, p 21). Weathering removes near-surface carbonates by dissolution and sulfide by oxidation. Weathered rock extends 6 to 12m below surfaces in the unglaciated portion of the Allegheny Plateau (Brady et al., 2000). Salt springs occur in Ohio Ecoregion 70, but they also occur in other states in Ecoregion 70.

Glaciation was evaluated as a possible natural cause of the apparent elevated background SC. There are two ecoregions in Ohio’s Allegheny Plateau, the Western Allegheny Plateau (Ecoregion 70) and the Lake Erie Drift Plain (Ecoregion 61). Ohio Ecoregion 70 is in the southwestern unglaciated portion of the larger Allegheny Plateau, and so dissolution is not expected to be greater than the rest of Ecoregion 70. Some headwaters of the Muskingum River located in the glaciated Ecoregion 61 drain into Ecoregion 70 and might naturally influence downstream SC in that drainage. However, the Ohio new area data provide empirical evidence that SC was not greater within the Muskingum drainage.

3.2.1.3. Physiography.: Ecoregion 70 has hilly and wooded terrain in both areas. The local relief of the unglaciated Allegheny Plateau in southeastern Ohio and western West

Virginia typically ranges from 61 to 229 m with peak elevations of around 610 m. Many of the rivers in this ecoregion are entrenched, as a result of the hilly terrain. The similar physiography in the two areas is expected to lead to similar surface and groundwater hydrology and the relatively uniform degree of deformation of the land should result in similar flow paths.

3.2.1.4. Geophysical and geochemical modeling.: Base-flow SC was estimated for Ecoregion 70 using a geochemical and geophysical model that predicts natural base-flow SC using geology, climate, soil, vegetation, topography, and other factors and is calibrated with references sites (Olson and Hawkins, 2012). Sources of data for the modeled estimates are listed in Appendix A.1. The estimated average stream baseflow SC was 180 μS/cm for the entire Ecoregion 70, 195 μS/cm for Ohio, 164 μS/cm for Pennsylvania, 176 μS/cm for West Virginia, and 172 μS/cm for Kentucky in Ecoregion 70. Fig. 3 shows the estimated SC for streams in Ecoregion 70. The model suggests that base-flow SC is slightly higher in Ohio (8.3% higher than the ecoregion average and 10.8% higher than West Virginia).

3.2.2. Summary of evidence from ecoregional characteristics—Natural dissolution of minerals and dilution in surface and groundwater are expected to be similar throughout Ecoregion 70, because similar climate, rock strata and physiography occur throughout the ecoregion. It should be noted that local anomalies occur but do not contradict the general uniformity of ecoregions identified by Fenneman and Johnson (1946). For example, whereas most fresh waters are dominated by Ca^{2+}, Mg^{2+}, HCO_3^-, and SO_4^- ions, there are natural salt (NaCl) springs in Central Appalachia.

The evidence based on ecoregional characteristics is summarized in Table 5. The ecoregional characteristics all provide evidence of similarity.

3.2.3. Evidence for consideration B: Empirically estimated background—If the SC background levels are similar in the two areas, the SC measurements should be similar. However, because of differences in sampling, the measurements are not equivalent. Therefore, various analyses and comparisons were performed.

Low SC values are background and should be similar. First, in the 4452 samples from 809 unique stream sites in the new area data set (Fig. 2), the minimum was 49 μS/cm compared to the original area which was 40 μS/cm. These similar minima are indicative of similar distributions at the low end and are evidence of similar background (Fig. 4). Second, background levels are often set at the 25th centile of regional data sets. This background estimate is higher for the new area (244 μS/cm) than for the original are (169 μS/cm). For highly disturbed regions, 10th centiles are used. This background estimate is also higher for the new area (163 μS/cm) than for the original are (120 μS/cm).

Background may be estimated more reliably from measurements at reference sites. The background level for the original area (210 μS/cm) is the upper 95% CL on the 75^{th} centile SC at reference sites. No equivalent value is available for the new area. That is because sites identified in the Ohio data sets as reference sites often had poor habitat scores and may have been designated as references for tiered usages associated with mine drainage. Therefore, samples from a set of minimally disturbed sites equivalent to the West Virginia reference set were not available. However, it was found, as expected, that a non-trivial proportion of sites in the new area (19%) have SC < 210 μS/cm.

It is also informative to consider monthly SC data for the two areas. Fig. 5 shows monthly box and whisker plots for the original and new areas. The lower quartiles (lower whiskers) overlap, as do the interquartile ranges (the boxes), for April through October, the months for which at least 30 sites were sampled in both areas. This evidence supports the similarity of background for most of the year.

These comparisons do not represent a census or proportional estimate of low SC streams in the new area, because the samples are not randomly selected. Rather, they are results of available measurements taken from two data sets developed for different applications in Ohio Ecoregion 70. The wadeable stream survey program's biological sampling protocol required that sites have a minimum depth of 10 cm in order to submerge Hester-Dendy substrates. In contrast, site selection of headwater streams in the new area and all streams in the original area did not have minimum depth specifications. As a result, the new area data set has more samples during lower flow from larger drainages (Tables 2 and 3). This may bias samples toward higher SC primarily due to increased potential anthropogenic inputs.

3.2.4. Evidence for consideration C: Distribution of low SC sites—If the sites with low SC are distributed across the new area, that suggests a broad distribution of sites representing background conditions. If, on the contrary, they existed only on a margin of the area that might suggest that they are associated with conditions in the adjoining area. If they occurred only in a few locations, then that would suggest that they are a result of local anomalies. In this example, low SC sites in the portion of Ecoregion 70 within Ohio show a similar distribution in all sampled subareas, except for the higher density of low SC to the southwest of the Scioto River (Fig. 2) located in the Shawnee State Forest. The spatial distribution of low SC sites is evidence of a low SC regional background.

3.2.5. Evidence for consideration D: Biological properties—The presence of obligate salt-intolerant species indicates that low SC streams occur in the sampled region, because they cannot survive without low SC. The loss of a taxon is characterized by the concentration at which it is extirpated. The extirpation concentration threshold is defined by the concentration below which 95% of the occurrences of the taxon are observed in a genus that declines to zero occurrences with increasing concentration (the XC_{95}) (U.S. EPA, 2011a; Cormier and Suter, 2013). The ionic niche inhabited by species can be narrow or wide and often extends above the background SC estimated at the 25th centile. So, the extirpation concentration (XC_{95}) for most species tends to be higher than background SC.

If salt-intolerant species have historically occurred in an area, the natural background must have been low enough to accommodate them. If salt-intolerant species currently occur in an area, the current background SC must be low enough for them to survive. Absence of salt intolerant species or genera alone is not definitive evidence that the regions are different, because historical records are often incomplete, other factors may limit occurrence, and current conditions may not allow recolonization. Furthermore, the SC tolerance of species varies within a genus and benthic invertebrate samples are often identified only to genus.

Sampling design and quality assurance (QA) should consider the comparability of field methods and taxonomic identification. In particular, sampling date is critical because the most salt-intolerant taxa are more likely to be collected in the first half of the calendar year. Sampling natural substrates is also necessary because many salt-intolerant taxa are shredders and are less likely to colonize deployed substrates.

In this example, invertebrate genera that require low SC waters occur in the new area (Ohio Ecoregion 70), and therefore, low SC conditions must occur in that region. Four of the seven genera in the original area data sets with XC_{95} values < 340 μS/cm, (*Alloperla, Diploperla, Heptagenia,* and *Ephemerella*) occur in the new area (OEPA, 2013). Seven genera in the original area with XC_{95} values <340 μS/cm that also occur in the combined data sets for West Virginia Ecoregions 69 and 70 occur in the new area (*Lepidostoma, Alloperla, Diploperla, Ephemerella, Clioperla, Heptagenia,* and *Nixe*) (OEPA, 2013; U.S. EPA, 2016a). Of the 139 genera included in the original area for which XC_{95} values are available, 90.6% are represented in the new area data set, indicating similarity of the biotas across state lines. The presence of genera that require low ion concentration levels indicates that low SC habitats occur in Ecoregion 70 in Ohio; however, the taxonomic identification was at the genus level and species and proportional representation of species may not be the same.

The similar natural background faunas, including salt-intolerant taxa, are evidence of similar background water quality.

3.2.6. Evidence for consideration E: Evidence relevance and reliability—

When determining similarity, relevance is reduced when there are differences between the sampling or measurement in the new area and in the original area. For example, sampling in the summer may bias the samples toward higher estimates of SC relative to sampling in other seasons. Sampling during precipitation events may bias toward lower SC estimates. Historical measurements represent conditions at that time and land uses and emissions may have changed (Schneider, 1965). Hence, when comparing two areas, the data should be derived under the same conditions in both places. For example, when data have not been collected in similar seasons, it may be necessary to normalize the data for Julian day. Most importantly, the samples should include minimally-disturbed or reference sites. Watersheds with >90% native vegetation are more likely to have low SC than areas that are developed. Even if >90% of an area is forested, proximity of a sampling site to a point source can have a greater effect than the proportion of the catchment in forest (Hopkins et al., 2013). Headwater streams are more likely to represent background SC, because there are likely to be fewer nonpoint and point sources.

In the characterization of background SC, sampling methods and designs can affect the reliability of results. Evidence of questionable data quality would suggest that the reported background may be inaccurate or biased (Section E of Table 1). The reliability of the SC measurements is evaluated by checking QA records for issues such as calibration of SC meters with solutions prepared for dilute waters, (i.e. low SC). In addition, data should be evaluated for apparent discrepancies. For example, if the lowest reported SC level is consistently the same round number (e.g. 100 μS/cm), this may be a reported detection limit and not a true measurement, and confidence in the measurement of background at low SC is lessened (Helsel, 2012). Conductivity meters should report temperature and be equipped with an automatic temperature compensation capability. Older equipment may have been used and measurements reported as electrical conductivity which may or may not have been normalized for temperature.

Natural SC regimes are the result of natural causes. Abrupt changes at political boundaries suggest that differences in results from the two regions are not natural. Such differences could be due to differences in sampling methods, sampling teams, land use practices, or other factors.

Results can be validated by independent estimates. Validation with respect to sampling variance can be performed by using an independent data set or reserving part of a large data set for validation. However, no additional reliable independent data sets were available and the new area data set was not large enough to partition. Alternatively, models that predict stream chemistry based on geology, climate, topography, and vegetation can provide an independent characterization and may even be useful for estimating natural background for smaller geographical areas (Olson and Hawkins, 2012).

In this example, the new area data set was used to assess SC regimes in Ohio. The reliability in terms of data quantity and QA procedures for data generation was judged to be high. The relevance in terms of similarity of sampling design was questionable. However, the combined wadeable-headwater new area data set contained at least some samples distributed through the year representing the full range of seasonal variation in SC (Table 2 and Fig. 5). The relevance was assessed as good with respect to seasonal distribution. Because Ohio's designated reference sites often had poor habitat, no reference-based estimate of background in the new area could be derived that would be comparable to the reference-based background in the original area.

Based on consideration of relevance and reliability described in Table 5, the suitability of SC measurements in the new area data set for characterizing background in Ohio Ecoregion 70 was judged to be ambiguous.

Biological data are subject to some of the same issues as SC data. Sensitive genera are less likely to be detected in the summer so they are less likely to be collected in an Ohio sample. Similarly, to the extent that Ohio data are more likely to represent disturbed areas, they are less likely to collect sensitive genera. However, these issues are less important for the biological evidence, which is based on occurrence in the region rather than abundance or frequency of occurrence. That is, there were enough spring samples in good quality Ohio streams to collect the sensitive genera. Hence, the biological data are suitable for the comparison.

Information concerning regional properties are straight-forwardly suitable. The geologic, physiographic, natural land cover, and climatic information is abundant, relevant and reliable.

3.3. Example case study: Synthesis

The weight of the entire body of evidence supports similarity of SC background in the two areas although the evidence was not strong. Measured stream SC in the new area (Ecoregion 70 in Ohio) was found to be higher than in the original area (Ecoregion 70 in West Virginia), but the relevance of the data is ambiguous, and other evidence supports similar background SC. The evidence collected and weighted in Table 5 is summarized in terms of the considerations, and an overall summary is presented for the hypothesis of sufficiently similar background in Table 6.

The climate, geology, and physiography are the same in the two areas, so the background SC is expected to be the same. A model based on geohydrology estimated that the mean SC in the new area is higher than the original area, but the two areas differed by only 10%. Hence, there is no evidence of a natural physical cause of a higher background for Ecoregion 70 in Ohio than in the rest of the ecoregion.

Because this example is a regional scale assessment, some substantially higher and lower local natural SC regimes may occur and require site-specific evaluation of applicability of a regionally derived benchmark or criterion, but that is true of all regional properties. Overall, a single background value would be practical for the region. If it appeared that certain watersheds or other subregions were substantially different, they could be delimited and the evidence that the apparent differences are natural could be reassessed.

4. Discussion

The concept of natural background is complex and has different interpretations in different circumstances. Background levels are conventionally defined by simple analyses of chemical concentrations (e.g. 25th centile, mean plus two sigma), but such analyses may not distinguish natural levels from best available levels and may be misleading if a consistent probability design is not used. Other evidence that should be considered include natural properties such as geology, climate, physiography and vegetation that determine natural levels; the occurrence of local sources; the spatial distributions of concentrations; and biological indicators of natural or altered conditions. No single quantitative method reliably analyzes those concepts or incorporates all relevant information, so the body of relevant evidence must be weighed. However, the assessment should not be ad hoc or informal (Cormier and Suter, 2008; U.S. EPA, 2016b). Rather, the assessment should follow an assessment framework containing planning, analysis, and synthesis and the process should assemble, weight, and weigh the relevant body of evidence. In particular, it should weight the evidence with respect to its relevance and reliability. We have illustrated the use of WoE within a condition assessment framework to determine the similarity between areas within a region of background SC. In this case, the other evidence collectively had more weight than the SC measurements. This WoE method might be used to determine if a benchmark, criterion, or restoration goal is appropriate to apply throughout an ecoregion or between two areas at any scale (U.S. EPA, 2016b; Suter et al., 2017).

The naturalness of apparent background is often ambiguous. Differences in estimated background may be unnatural in two ways. (1) It may be unnatural in the sense that it is not real. It may be an artifact because the sample size is small and the estimate is biased by chance. In addition, differences in sampling design including location and seasonality and differences in measurement techniques may bias results. (2) A difference may also be unnatural because it is altered from natural conditions by human activities. The obvious sources of unnatural SC are human disturbances of the land (e.g. mining, construction, inadequate restoration) and effluent releases of dissolved ions to surface waters (e.g. industrial, municipal or domestic waste). The WoE method can address both causes of unnatural differences, whereas a simple analysis of concentration data cannot.

The types of evidence for assessing background will vary with the application and characteristics of the areas being assessed (U.S. EPA, 2011b). The WoE process begins by assembling the relevant evidence which requires thinking broadly about relevance. Four major considerations for the case were identified, which may be applicable to other assessments of background. (1) What natural or anthropogenic properties of the area determine the observed levels? (2) What are the measured levels and what background levels do they suggest? (3) Does the spatial distribution of measured levels suggest that natural or anthropogenic factors influence background? (4) Is the biota consistent with natural background? Each of those considerations suggests multiple pieces of evidence that may be generated from available data. Each of those pieces is weighted (Table 5) and the weights are combined into a weight for each consideration (Table 6).

Different degrees of variation in background may be relevant to the application. In this case, the background was assessed to evaluate whether an example criterion developed in one area could be used throughout the entire ecoregion. No two areas are identical, so it is necessary to decide what difference is important enough to suggest separation and which estimated differences are relevant and reliable. The minima in the two areas differ by 9 µS/cm (22%). The 25th centiles of all sites in each area (a conventional estimate of background for areas that are somewhat modified) differ by 75 µS/cm (31%). The 10^{th} centiles of all sites in Ohio (a conventional estimate of background for areas that are highly modified) was 6 µS/cm less than in the potentially less disturbed original area (3.6%). The results of the geophysical models of the two areas differ by 19 µS/cm (10%). Most of these estimates indicate a higher background in Ohio, but the absolute and relative differences vary greatly among estimates. The geophysical model, which includes only natural variables and is not influenced by differences in development or sampling, suggests a small difference. The 25th centile levels suggest a larger difference but is influenced by unnatural differences and sampling bias. The other evidence would seem to support little difference in SC which is consistent with the modeled results. There is no difference in regional properties that would cause a large difference in SC. The low SC levels in the new area are distributed as if they were from a residue of natural streams. And, the occurrence of sensitive invertebrates suggests that the measured low SC levels are really natural background.

The decision to assign a background value to an area may depend on more than weighing the evidence concerning natural background levels. This is a policy decision that hinges upon societal values and on the change in the effect relative to the change in the stressor. One consideration in addition to the WoE is the consequences of assigning a high background level that is not protective. Using a background-to-criterion model and flowchart method

(U.S. EPA, 2016a; Cormier et al., 2018a,b) and background of 244 µS/cm in the new area (the 25th centile), 5% of genera would be extirpated at 439 µS/cm with a lower 50% prediction limit of 366 µS/cm. Using the background-to-criterion model and the background estimated for the new area by the geochemical model (195 µS/cm), 5% of genera would be extirpated at 378 µS/cm with a lower 50% prediction limit of 315 µS/cm. Professional judgment guided by a weight of evidence is needed. Additional analyses can be performed to determine consequences of choosing one or the other estimate. For example, if the example criterion of 340 µS/cm is the true value for 5% extirpation, one might ask what is the proportion of genera lost at 366 µS/cm, the lower 50% prediction limit as recommended using the background-to criterion model (Cormier et al. 2018b)?

An additional consideration is the types of uncertainties that are acceptable. If the new area is judged to be different, the SC criterion would be estimated with an uncertain estimate of background input to a background-to-criterion regression model (Cormier et al., 2018a). Despite the uncertainty, this option may seem preferable because the criterion would rely upon samples taken in the new area. If the original and new areas were judged to be the same, the new area criterion would be based on a relationship between SC and biota in a large set of data from the original area. This option uses a high-quality relationship and would bypass the questionable SC data and the background-to-criterion regression model, which eliminates sources of uncertainty. However, that option may seem undesirable, because it uses the relationship between SC and biota from many other areas and not the new area.

A third additional consideration is the acceptable level of variation in natural SC. That is, is the background across the ecoregion uniform enough to apply a single criterion? The modeled mean base flow varies by >200 µS/cm in 20% of stream segments in about a third of ecoregions in the USA (Cormier et al., 2018b). Therefore, there is a need for professional judgment when applying a single value. There are several options that might be considered. One might consider using some sort of a tiered approach or developing site or resource specific criteria.

In either option, areas with different backgrounds will need to be delineated based on modeled and measured attributes. Also, anomalies will arise and natural versus anthropogenic background may need to be distinguished. Here is where a formal weight of evidence can be particularly helpful.

A key benefit of using a WoE approach to assess background or other field conditions is that it requires deliberate and conscientious considerations of many factors that may affect use of a background level in an area. Consequently, the assessors understand the situation more deeply than if a simple statistical criterion was used and that understanding can support a well-informed conclusion. However, this benefit implies the disadvantage of requiring expertise and effort. A related benefit of a formal weighing of evidence is that the rationale is transparent and can be updated as new information becomes available. The disadvantage of this approach is that documenting and justifying assumptions and inferences may slow decision making. When weighing evidence for a more complex case, it would be appropriate to make the process even more formal and transparent by scoring each piece of evidence with respect to relevance, reliability, and strength (U.S. EPA, 2016b; Suter et al., 2017). In sum, weighing evidence of sufficient similarity is possible when it is not possible to collect and analyze hundreds of samples in both areas using the same protocol. Even when such sampling is performed, weighing other evidence can support findings or explain differences.

5. Conclusions

An assessment of the similarity of SC background in Ecoregion 70 in Ohio and West Virginia concluded that, based on the WoE, it is likely that background is similar across the ecoregion. Conventional estimates of background are not similar, but they are influenced by differences in sampling design, in seasonal distribution, in degree of disturbance, and other factors. Ideally, a sampling program focusing on undisturbed watersheds and designed to determine background would be performed. This case illustrates the use of a formal assessment framework and inference by weighing the body of evidence when making judgments concerning background. This result is significant because the concept of background is important in environmental assessment, and because no one analytical method has been found to reliably characterize background.

6. Data sets

Data are contained in two files (Cormier, 2017). Data sets are available at the U.S. EPA Environmental Dataset Gateway (https://doi.org/10.23719/14024.18). Data contains sampling station locations with physical and chemical data. Data: stations 508.xlsx (Ohio data set), env.bio70508.xlsx (WV biological station dataset). Links to the data for the base-flow predictive model is available in Appendix Table A.1. Table A1 shows the data sources for modeled base-flow conductivity used to develop Fig. 3. Empirically modeled base-flow SC for streams in Pennsylvania, Ohio, West Virginia, and Kentucky.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors have no conflict of interest to report. This work was supported by and prepared at the EPA, National Center for Environmental Assessment (NCEA), Cincinnati Division, and Office of Water, Health and Ecological Criteria Division, Washington, DC. TetraTech prepared data sets and provided statistical analyses under U.S. EPA contract #EP-C-12– 060. The authors are indebted to the work of field and laboratory personnel who collected and analyzed the primary data. As part of external review of “Draft: Field-based methods for Developing Aquatic Life Criteria for Specific Conductivity,” the research presented here was reviewed by five scientists in an independent contractor-lead peer review. The manuscript has been subjected to the agency’s peer and administrative review and approved for publication. However, the views expressed are those of the authors and do not necessarily represent the views or policies of the EPA. Kari Ohnmeis of Highlight Technologies, LLC and Linda Tackett of CACI International Inc., edited and formatted the document. Charlotte Moreno assisted with on-line preparation. Constructive comments from Dan Petersen, Joseph Flotemersch, and anonymous reviewers helped to substantially improve this manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Abbreviations:

C	Carbon
Ca	Calcium
Cl	Chloride
H	Hydrogen
K	Potassium
km	Kilometer
Max	Maximum
Mg	Magnesium
mg/L	Milligrams per liter
Min	Minimum
Na	Sodium
O	Oxygen
Ohio EPA	Ohio Environmental Protection Agency
pH	Potential of hydrogen
QA	Quality assurance
RBP	Rapid Bio-assessment Protocol
SC	Specific conductivity
SU	Standard units
US EPA	United States Environmental Protection Agency
WoE	Weight of evidence
XC95	Extirpation concentration values at the 95th centile
XCD	5th centile of XCD
µS/cm	Micro Siemens per centimeter

References

Anning DW, Flynn ME, 2014 Dissolved-Solids Sources, Loads, Yields, and Concentrations in Streams of the Conterminous United States. U.S. Geological Survey, Reston, VA.

Barbour MT, Gerritsen J, Snyder BD, Stribling JB, 1999 Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates, and Fish. 2nd ed. U.S. Environmental Protection Agency, Office of Water, Washington, DC (EPA/841/B 99/002). https://www3.epa.gov/region1/npdes/merrimackstation/pdfs/ar/AR-1164.pdf.

Brady KBC, 1998 Natural groundwater quality from unmined areas as a mine drainage quality prediction tool In: Brady KBC, Smith MW, Schueck J (Eds.), Coal Mine Drainage Prediction and Pollution Prevention in Pennsylvania. PA Department of Environmental Protection, Harrisburg, PA https://www.osmre.gov/resources/library/pub/cmdppp.pdf.

Brady K, Hornberger R, Chisholm W, Sames G, 2000 Chapter 2: how geology affects =mine drainage prediction In: Kleinmann R (Ed.), Prediction of Water Quality at Surface Coal Mines. West Virginia University National Mine Land Reclamation Center, Morgantown, WV.

Cormier SM, 2017 Data for Assessing Background Levels of Specific Conductivity Using Weight of Evidence 508 Compliant. (dataset). 10.23719/1402418.

Cormier SM, Suter G, 2008 A framework for fully integrating environmental assessment. Environ. Manag. 42 (4), 543–556.

Cormier SM, Suter GW, 2013 A method for deriving water-quality benchmarks using field data. Environ. Toxicol. Chem. 32 (2), 255–262. [PubMed: 23147651]

Cormier S, Zheng L, Flaherty CM, 2018a A field-based model of the relationship between extirpation of salt-intolerant benthic invertebrates and background conductivity. Sci. Total Environ. 10.1016/j.scitotenv.2018.02.044.

Cormier S, Zheng L, Hill R, Novak R, Flaherty CM, 2018b A flow-chart for developing water quality criteria from two field-based methods. Sci. Total Environ. 10.1016/j.scitotenv.2018.01.137.

Daly C, Neilson R, Phillips D, 1994 A statistical topographic model for mapping Climatological precipitation over mountainous terrain. J. Appl. Meteorol. 33, 10–158.

European Union, 2000 Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for community action in the field of water policy. Off. J. Eur. Communities 27, 1–72.

European Union, 2006 Directive 2006/118/EC of the European Parliament and of the Council of 12 December 2006 on the protection of groundwater against pollution and deterioration. Off. J. Eur. Union 372:19–31. http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:372:0019:0031:EN:PDF

Fenneman N, Johnson D, 1946 Physiographic Divisions of the Conterminous U.S. U.S. Geological Survey, Reston, VA.

Galuszka A, 2007 A review of geochemical background concepts and an example using data from Poland. Environ. Geol. 52, 861–870.

Gough L, Crock J, 1997 Distinguishing Between Natural Geologic and Anthropogenic Trace Element Sources. 1105. Denali National Park and Preserve. U.S. Geol. Surv. Circ., pp. 57–71.

Griffith MB, 2014 Natural variation and current reference for specific conductivity and major ions inwadeable streams of the coterminous U.S. Freshw. Sci. 33 (1), 1–17.

Helgen S, Moore J, 1995 Natural background determination and impact quantification in trace metal-contaminated river sediments. Environ. Sci. Technol. 30, 129–135.

Helsel D, 2012 Statistics for Censored Environmental Data Using Minitab and R. 2nd ed. John Wiley and Sons, Inc, Hoboken, NJ.

Hem J, 1985 Study and Interpretation of the Chemical Characteristics of Natural Waters USGS Water Supply Paper 2254. Department of the Interior, U.S. Geological Survey, Washington, DC https://pubs.usgs.gov/wsp/wsp2254/pdf/wsp2254a.pdf.

Hinsby K, de Melo M, Dahl M, 2008 European case studies supporting the derivation of natural background levels and groundwater threshold values for the protection of dependent ecosystems and human health. Sci. Total Environ. 401, 1–20. [PubMed: 18486193]

Hopkins R, Altier B, Haselman D, Merry A, White J, 2013 Exploring the legacy effects of surface coal mining on stream chemistry. Hydrobiologia 713, 87–95.

Luo X-S, Xue Y, Wang Y-L, Cang L, Xu B, Ding J, 2015 Source identification and apportionment of heavy metals in urban soil profiles. Chemosphere 127, 152–157. [PubMed: 25698100]

Mast MA, Verplanck PL, Wright WG, Bove DJ, 2007 Characterization of background water quality. Professional Paper 1651. Chapter E7 In: Church SE, von Guerard P, Finger SE (Eds.), Integrated Investigations of Environmental Effects of Historical Mining in the Animas River Watershed, San Juan, County, Colorado. U.S. Geological Survey,Washington,DC https://pubs.usgs.gov/pp/1651/downloads/Vol1combinedChapters/vol1_chapE7.pdf.

Molinari A, Chidichimo F, Straface S, Guadagmini A, 2014 Assesment of natural background levels in potentially contaminated coastal aquifers. Sci. Total Environ. 476, 38–48. [PubMed: 24448030]

ODGS, 2006 Bedrock Geologic Map of Ohio: Map BG-1, Generalized Page-Size Version with Text, 2p, Scale 1:2,000,000. Ohio Department of Geologic Survey, Department of Natural Resources, Columbus, OH http://geosurvey.ohiodnr.gov/portals/geosurvey/PDFs/BedrockGeology/BG-1_8.5×11.pdf.

OEPA, 2013 Ohio EPA Macroinvertebrate Taxa List. Ohio Environmental Protection Agency, Department of Surface Water, Columbus, OH http://epa.ohio.gov/Portals/35/documents/Macro_Taxa_List.pdf.

Olson J, Hawkins C, 2012 Predicting natural base-flow stream water chemistry in the western United States. Water Resour. Res. 48 (2) (W02504).

Omernik JM, 1987 Ecoregions of the conterminous United States. Ann. Assoc. Am. Geogr. 77 (1), 118–125.

Page DS, Boehm PD, Douglas GS, Bence AE, Burns WA, Mankiewicz PJ, 1996 The natural petroleum hydrocarbon background in subtidal sediments of Prince William Sound, Alaska, USA. Environ. Toxicol. Chem. 15, 1266–1281.

R Development Core Team, 2011 R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria http://www.R-project.org.

Reimann C, Garrett R, 2005 Geochemical background—concept and reality. Sci. TotalEnviron. 350,12–27.

Reimann C, Filzmoser P, Garrett RG, 2005 Background and threshold: critical comparison of methods and determination. Sci. Total Environ. 346, 1–16. [PubMed: 15993678]

Runnells D, Shepherd T, Angino E, 1992 Metals in water, determining natural Background concentrations in mineralized areas. Environ. Sci. Technol. 26 (12), 2316–2323.

Schneider W, 1965 Water resources of the Appalachian region, Pennsylvania to Alabama USGS Hydrologic Atlas 198. U.S. Department of the Interior, U.S. Geological Survey, Washington, DC https://pubs.er.usgs.gov/publication/ha198.

Schruben P, Arndt R, Bawiec W, 1997 Geology of the Conterminous United States at 1: 250,000 Scale A Digital Representation of the 1974 P.B. King and H.M. Beikman Map: U.S. Geological Survey Digital Data Series DDS-11, Release 2. U.S. Department of the Interior, U.S. Geological Survey, Washington, DC.

Smith R, Alexander R, Schwarz G, 2003 Natural background concentrations of nutrients in streams and rivers of the conterminous United States. Environ. Sci. Technol. 37, 3039–3047. [PubMed: 12901648]

Stoddard JL, Larsen DP, Hawkins CP, Johnson RK, Norris RH, 2006 Setting expectations for the ecological condition of streams: the concept of reference condition. Ecol. Appl. 16 (4), 1267–1276. [PubMed: 16937796]

Sucharova J, Suchara I, Hola M, Reimann C, 2014 Contemporary lead concentration and stable isotope ratio distribution in forest moss across the Czech Republic. Appl. Geochem. 40, 51–60.

Suter G, Cormier S, Barron M, 2017 A weight-of-evidence framework for environmental assessments: inferring qualities. Integr. Environ. Assess. Manag. 13 (6): 1038–1044.10.1002/ieam.1954. [PubMed: 28613433]

U.S. EPA, 2000 Nutrient Criteria: Technical Guidance Manual; Rivers and Streams. U.S. Environmental Protection Agency, Office of Water, Washington, DC (EPA 822/B00/002). https://www.epa.gov/nutrient-policy-data/criteria-development-guidance-rivers-and-streams.

U.S. EPA, 2002 Guidance for Comparing Background and Chemical Concentrations in Soil for CERCLA Sites. Washington, DC: U.S. Environmental Protection Agency, Office of Emergency and Remedial Response, Washington, DC (EPA 540/R-01/003). http://dec.alaska.gov/spar/csp/guidance_forms/docs/background.pdf, https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=9101ALZR.TXT

U.S. EPA, 2011a A Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams. U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, Washington, DC (EPA/600/R-10/023F). http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=23=3809.

EPA Author Manuscript

EPA Author Manuscript

U.S. EPA, 2011b Landscape and Predictive Tools: A Guide to Spatial Analysis for Environmental Assessment. U.S. Environmental Protection Agency, Risk Assessment Forum Washington, DC https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P100QIFN.TXT.

U.S. EPA, 2016a Public Review Draft: Field-Based Methods for Developing Aquatic Life Criteria for Specific Conductivity. U.S. Environmental Protection Agency, Office of Water, Washington, DC (EPA-822-R-07–010). https:/www.epa.gov/wqc/draft-fieldbased-methods-developing-aquatic-life-criteria-specific-conductivity.

U.S. EPA, 2016bWeight of Evidence in Ecological Assessment. U.S. Environmental Protecion Agency, Risk Assessment Forum, Washington, DC (EPA/100/R-16/001). https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=335523.

Van de Meent D, Aldenberg T, Canton J, van Gestel C, Sloof W, 1990 Desire for Levels Background Study for the Policy Document “Setting Environmental Quality Standards for Water and Soil”. RIVM, Bilthoven, the Netherlands (Report Number 670101 002).

Woods A, Omernik J, Brown D, Kiilsgaard C, 1996 Level III and IV Ecoregions of Pennsylvania and the Blue Ridge Mountains, the Ridge and Valley, and Central Appalachians of Virginia, West Virginia, and Maryland. U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Corvallis, OR (EPA/600/R-96/077).

Woods A, Omernik J, Brown D, 1999 Level III and IV Ecoregions of Delaware, Maryland, Pennsylvania, Virginia, and West Virginia. U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Corvallis, OR https://nctc.fws.gov/courses/csp/csp3200/resources/documents/epa_region_3_eco_desc.pdf.

WVGES, 2011 Geologic Map of West Virginia. Map 25AWest Virginia Geological and Economic Survey, Morgantown, WV http://www.wvgs.wvnet.edu/www/maps/Geologic_Map_of_West_Virgini-Map25A_lowres.pdf.

Zhao F, McGrath S, Merrington G, 2006 Estimates of ambient background concentrations of trace metals in soils for risk assessment. Environ. Pollut. 148, 221–229.

EPA Author Manuscript

EPA Author Manuscript

Highlights

- Water quality criteria for dissolved ions are related to background levels.
- If ecoregions are sufficiently similar, background levels should be the same.
- Evidence of similar background was weighed in a formal assessment process.
- A case study showed similar background specific conductivity within an ecoregion.
- Background is best characterized by weighing the body of relevant evidence.

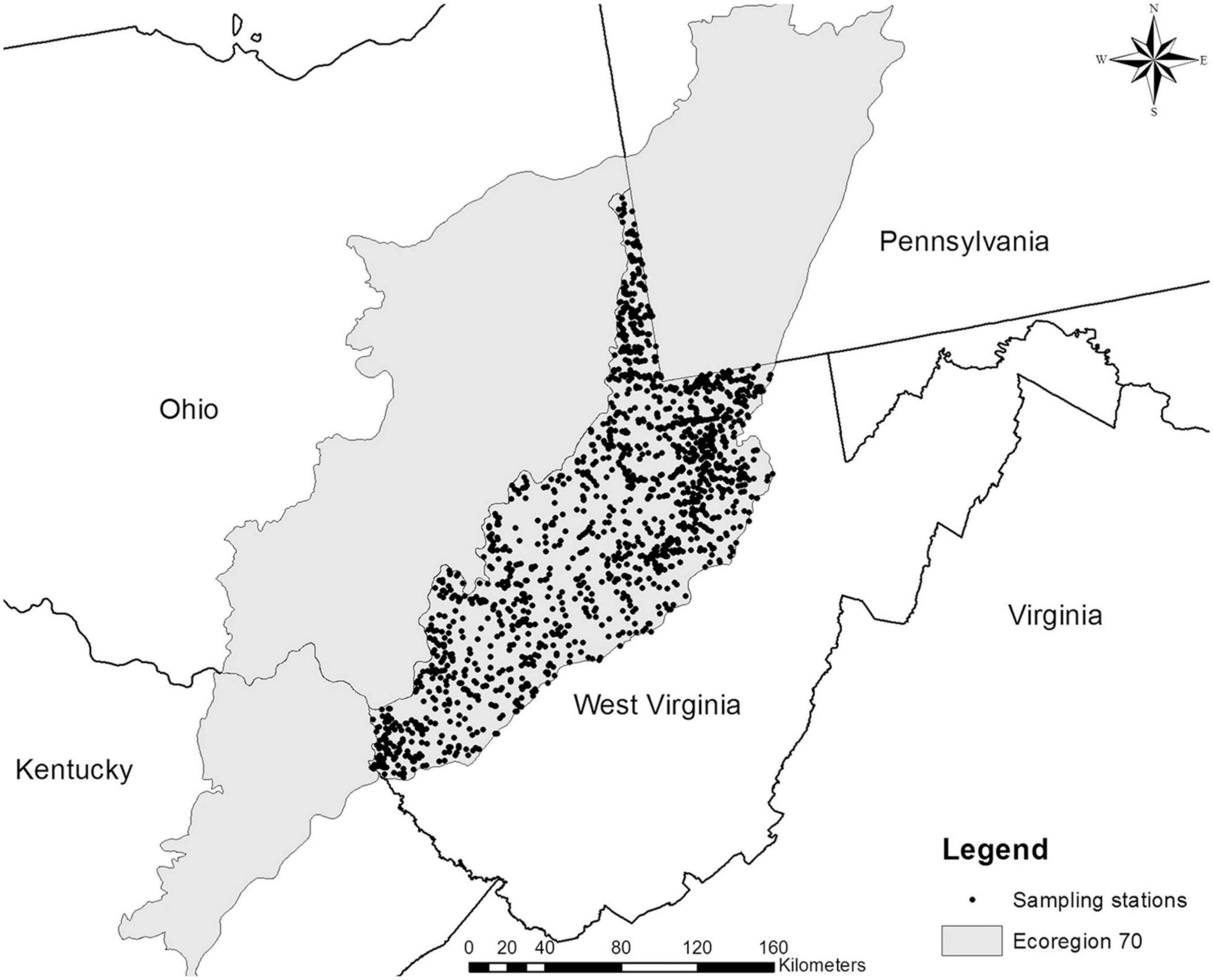


Fig. 1.
Ecoregion 70 extends from the northeast corner of Kentucky through Ohio and West Virginia in southwest Pennsylvania. Sampling sites (stations) (N = 1695) in the data set that were used to estimate background are indicated as points. The grey area in Ohio is the focus of this assessment. Source: U.S. EPA, 2016.

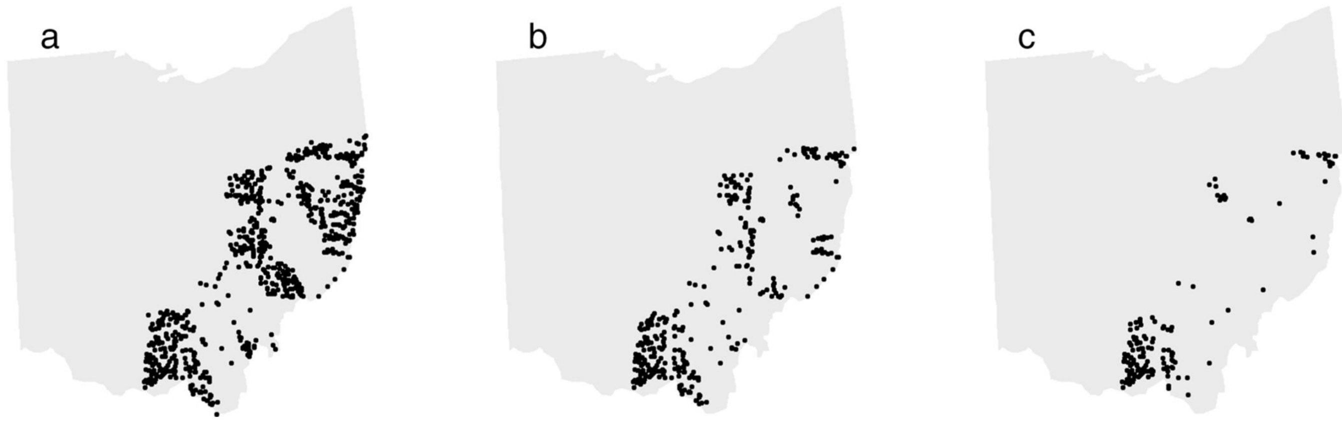


Fig. 2.

Distribution of sampled sites in the new area data set. Low SC sites are representatively dispersed throughout the sampled areas. (a) All sites from the new area data set ($N = 809$), black points roughly delineate Ecoregion 70 on the grey Ohio map (b) samples <340 μS/cm, the example ecoregional benchmark, ($N = 353$), and (c) all samples less than background (210 μS/cm) for the original area data set ($n = 154$). More low SC sites occur west of the Scioto River in the lower left. Source: U.S. EPA, 2016.

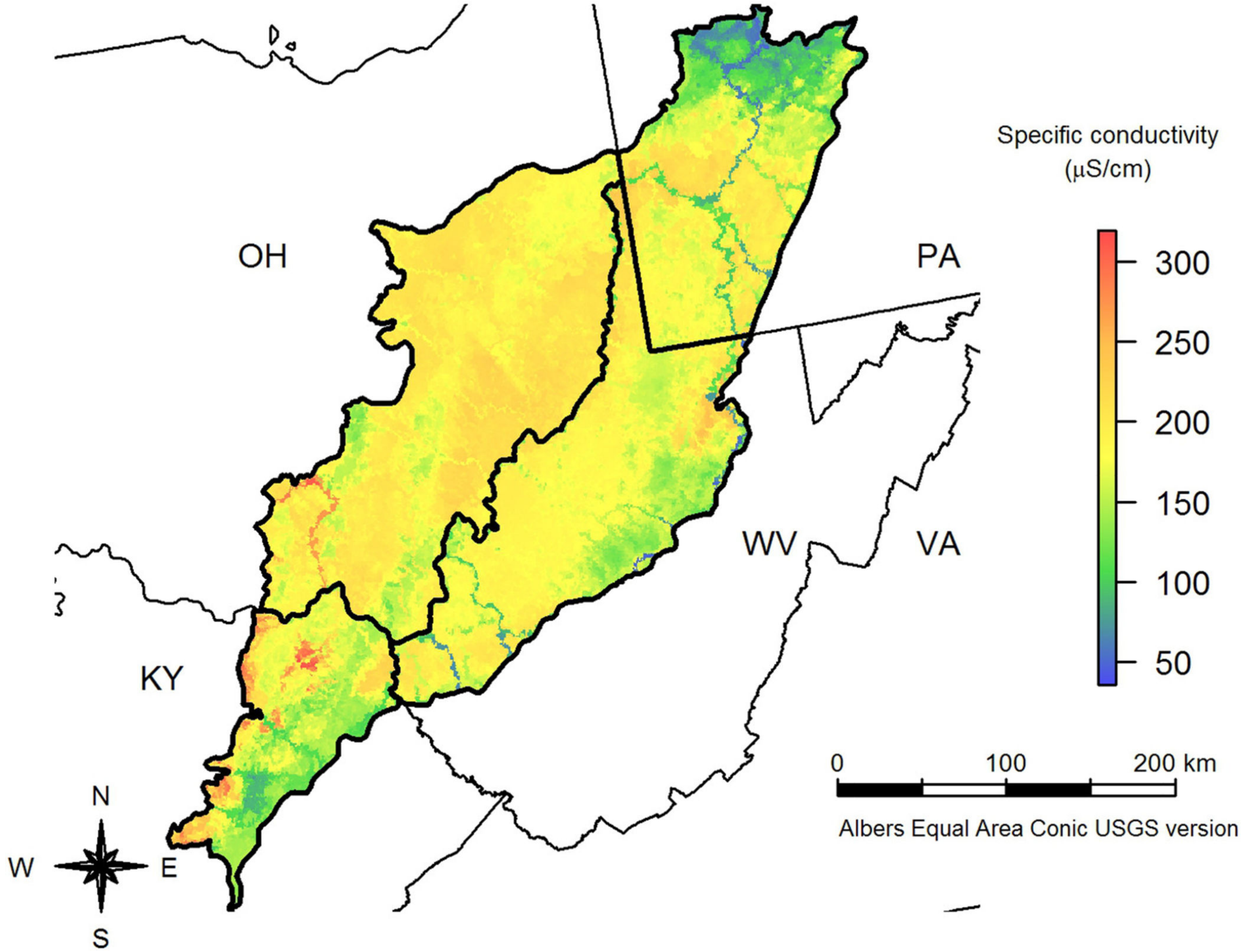


Fig. 3.
Geophysically modeled base-flow SC for streams in Pennsylvania, Ohio, West Virginia, and Kentucky. The SC range in the yellow, central portion of the ecoregion is expected to be between 161 and 230 μS/cm. Source: U.S. EPA 2016, (data sources: Appendix A.1).

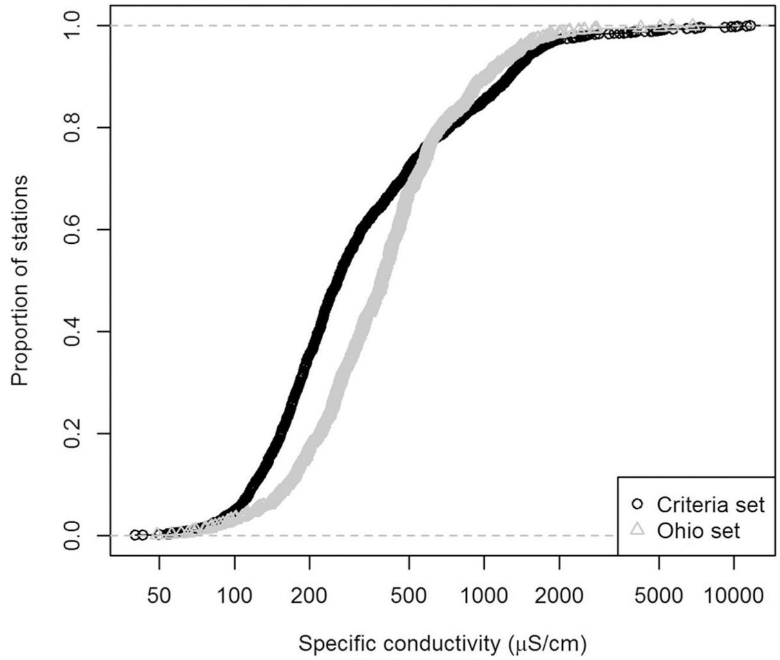


Fig. 4.
Cumulative distribution function of SC measurements from the original area data set and the new area data set in Ecoregion 70. At the lowest background SC (<100 μS/cm), the data distributions overlap, but Ohio levels are higher at intermediate proportions. 25th centiles are 169 and 244 μS/cm, but data sets are not directly comparable. Source: U.S. EPA (2016).

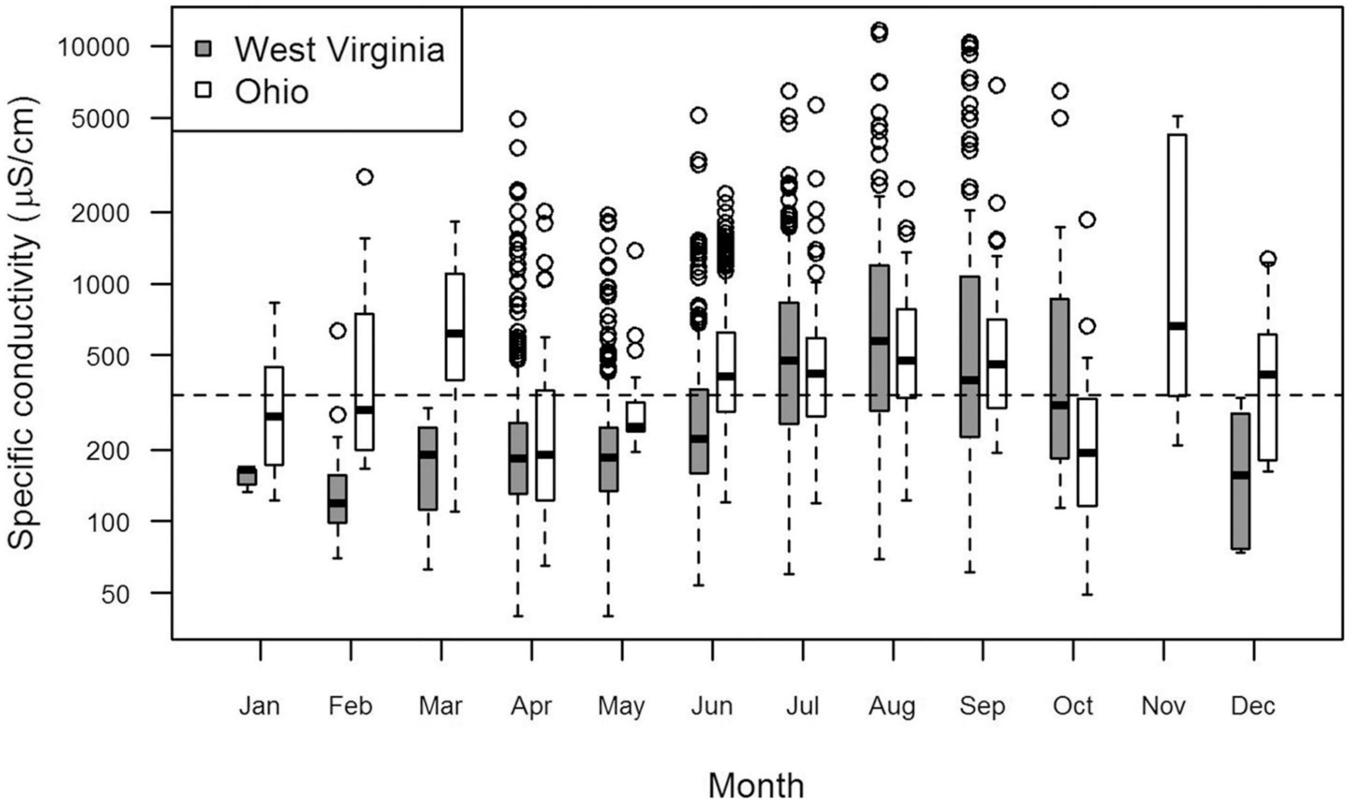


Fig. 5.
Specific conductivity distributions from Ecoregion 70 in the original area (West Virginia) (filled) and the new area (Ohio) (unfilled). The ranges of the lowest quartile SC values (lower vertical whiskers) overlap for months when >30 unique sites were sampled (April–October). The horizontal line is the example chronic criterion for the original area. Source: U.S. EPA (2016).

EPA Author Manuscript

EPA Author Manuscript

Table 1

Evidence for characterizing estimated background SC and for assessing its departure from natural background SC regimes, organized by considerations (A−E) and types of evidence (1−27).

	Considerations and types of evidence	Rationale
A.	Regional properties	
1	Similar climate across the region.	Ecoregions are classified based on shared climate, so differences within an ecoregion are less likely than among ecoregions. However, similarity must be confirmed. Climate affects vegetative cover and precipitation, which could affect background levels.
2	Similar geology across the region.	Ecoregions are classified based on shared geology, so differences within an ecoregion are minimized, and differences among ecoregions are more likely. However, similarity must be confirmed. Geology and soils can affect the background level.
3	Similar physiography across the region.	Ecoregions are classified based on shared physiography, so differences within an ecoregion are minimized relative to differences among ecoregions. However, similarity must be confirmed. Physiography can affect the hydrology and interaction time of water and minerals.
4	Geophysical modeling estimates low variance across the region.	Models that predict stream chemistry based on geology, climate, topography, and vegetation do not include human additions, so they should estimate background. Some models are available (e.g. Olson and Hawkins 2012).
B.	Empirically estimated background	
5	Lowest observed SC in the new area is approximately as low as in the original area.	Low SC is indicative of a minimally affected (natural background) condition, but individual low SC observations may be an artifact of sampling during or immediately after a rainstorm.
6	Similar 10^{th} and 25^{th} centiles of the full data sets in both areas or similar distributions.	Similar SC at centiles that may be considered background limits or similar SC distributions are evidence that background SC is the same.
7	The proportion of measured sites less than the background level in the original region (210μS/cm) is not negligible.	The greater the proportion of sites below the estimate of background level (210μS/cm for the case example), the more likely that the background is not different.
8	The 75^{th} centiles of reference sites are similar in the two areas.	Reference sites that are undisturbed should have low SC that is close to natural background. However, when anthropogenic changes are widespread, best available or least-disturbed sites may not accurately reflect natural background. In such cases, a description of the SC distribution may be more informative than classification as reference.
9	The lower quartile of the new area overlaps the lower quartile of the original area in most months.	When the lower quartiles consistently overlap during different time periods, the overall background also overlaps. The wider the overlap, the more likely that the background is not different.
C.	Distribution of low SC sites	
10	Low SC sites (<210 μS/cm) are spatially dispersed.	Sites with SC as low as background in the original area are dispersed throughout the new area suggesting that background is similar across the region.
11	Sites with SC less than a SC benchmark are spatially dispersed.	Moderately low SC sites dispersed throughout the new area suggest that a low natural background is representative of the region. In this example, the proposed example criterion limit was used to assess spatial uniformity of moderately low background levels.
D.	Biological indicators of background	
12	Salt-intolerant genera (XC_{95}<340μS/cm) are found in the historical record.	Historically, species adapted to the natural SC regime should have occurred in the region. The historical presence of salt-intolerant genera in the compared areas is indicative of low SC systems in the past.
13	Salt-intolerant genera (XC_{95}<340μS/cm) are found in recent records.	The presence of salt-intolerant genera is indicative of low SC systems in the region. Sampling date is critical. In the study area, most salt-intolerant taxa are more likely to be collected in the first half of the calendar year. Sampling natural substrates is necessary because many salt-intolerant taxa are shredders and grazers.
14	Regional biota is similar.	If the genera are similar between the two areas, the habitats are expected to be similar. However, salt-intolerance of species can differ widely within a genus. Geographically dispersed ecoregions with similar background may have different species, but both have a cadre of similarly salt-intolerant taxa.

	Considerations and types of evidence	Rationale
E.	Relevance and reliability of evidence	
15	Larger data sets are more reliable.	Large data sets provide more confidence in the evidence derived from them.
16	Quality assured instream chemical measurements are more reliable.	Meter calibration should be performed using appropriate standards. Otherwise, the measurements may be inaccurate.
17	Appropriate reporting units are required.	Units should be reported as SC. Otherwise, direct comparison of the measurements is inappropriate.
18	Comparisons are appropriate only if ionic mixtures are similar.	Data sets must be checked to eliminate sites with dissimilar ionic mixtures. Otherwise, the background estimates in the new area may not be comparable or reliable. Dissimilar mixtures are also suggestive of anthropogenic sources.
19	Comparisons are more reliable if collection methods and sampling windows are similar.	Data should be collected in similar seasons or be normalized for Julian day. Biological sampling methods should be similar. Otherwise, the estimates may not be comparable or reliable.
20	Validation increases reliability.	If more than one data set is available or another method is available for analysis and they yield similar results, then the results are validated and are less likely to be due to chance.
21	Inclusion of likely background sites makes identification of background SC more probable.	Inclusion of headwater (low Stahler order) streams in the sample increases the likelihood that background SC is represented, because they tend to receive fewer anthropogenic emissions. Watersheds with native vegetation are also more likely to have background water quality. The proximity and proportion of point sources or anthropogenically altered subsurface or land cover near the sampling sites decreases the likelihood that a site represents background. Even if >90% of area is forested, proximity of a sampling site to a point source has a greater effect than the proportion of vegetated area in the catchment (Hopkins et al. 2013).
22	Abrupt changes in SC values at political boundaries makes it likely that differences are not natural.	Abrupt changes in SC at political boundaries suggest that differences in SC are due to differences in methods or anthropogenic sources/practices and not natural causes. Such changes do not represent natural differences and support the argument that apparent differences in background SC between two areas are artifacts.
23	Reference site quality independently verified.	Reference sites are often considered background, but they may not be. Because identification of reference sites may be subjective, independent verification or comparison to another reference data set is desirable. References sites may represent the best of a class, e.g. best mined areas.
24	Reference sites are selected in part based on presence of high biological diversity.	Biological conditions for reference sites are important for verification of reference status.
25	Reference sites have high quality instream and riparian habitat (e.g. RBP score >140; Barbour et al. 1999).	Unless habitat quality is high, sites are unlikely to represent true reference water quality.
26	At reference sites, water quality parameters other than SC should reflect minimally disturbed conditions.	Unless all water quality parameters reflect minimally affected conditions, sites are unlikely to represent background.
27	Reference sites have minimal human disturbance of geological parent material.	Geological disturbances such as quarries and road cuts are likely to raise SC levels.

SC = specific conductivity; XC = extirpation concentration; RBP = Rapid Bio-assessment Protocol.

Table 2

Number of samples and sites from the new area data set with reported percentage meeting data acceptance criteria for specific conductivity (SC). Number of samples is shown for each month.

	Month												
Number of	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**	**Total**
Samples	101	118	151	132	167	666	890	957	643	380	119	128	4452
Unique sites	14	22	27	35	21	248	181	90	46	103	6	16	809
Percentage of total new area	1.7	2.7	3.3	4.3	2.6	30.7	22.4	11.1	5.7	12.7	0.7	2.0	100
Percentage of total original area	0.2	1.6	0.5	17.4	18.4	14	16	21.1	9.3	1.2	0	0.2	100

Table 3

Catchment sizes represented in the new area data set.

Proportion represented in the data set based on centile	Catchment original area (km^2)	Catchment new area (km^2)
Minimum	0.17	0.26
25^{th} centile	2.88	13.7
50^{th} centile	9.1	32.1
75^{th} centile	38.2	158
Maximum	3912.2	20850[a]

[a]Muskingum River

Table 4

Water chemistry measurements from new area data set, Ohio Ecoregion 70.

SC or ion	*N*	Min	10th	25th	50th	75th	Max
		Centile					
SC (μS/cm)	809	49	163	244	387	586	6838
HCO_3^- (mg/L)	550	3.05	34.37	69.98	114.1	179.3	380.6
SO_4^{2-} (mg/L)	560	5.2	23.7	35.33	61.1	141	1720
Cl^- (mg/L)	601	2.5	2.5	7.9	14.1	26.3	259
Ca^{2+} (mg/L)	588	6	16	27	45	71	419
Mg^{2+} (mg/L)	586	2	7	9	14	21	218
Na^+ (mg/L)	588	2.5	5	8	12	24	328
K^+ (mg/L)	588	1	2	2	3	4	29
pH (SU)	809	6.05	7.26	7.61	7.85	8.13	9.18
[a]$(HCO_3^- + SO_4^{2-})/Cl^-$	550	1.52	5.18	7.99	14.12	26.92	225.89

[a]Value within category calculated from individual sample ion concentrations in mg/L.

Min = min; Max = maximum; SC = specific conductivity, mg/L = milligrams per liter; H = hydrogen, C = carbon; O = oxygen; Cl = chloride; Ca = calcium; Mg = magnesium; Na = sodium; K = potassium; pH = potential of hydrogen; SU = standard units

Table 5

Evidence organized by consideration for characterizing current background specific conductivity (SC) in the new area (Ohio Ecoregion 70) and for assessing its similarity to the original area (West Virginia Ecoregion 70)[a].

	Types of evidence	Results
A.	Regional properties	
1	Similar climate across the region.	Climate is similar across the ecoregion. (+)
2	Similar geology across the region.	Geological strata are similar across the ecoregion. (++)
3	Similar physiography across the region.	The entire ecoregion has a hilly physiography. (+)
4	Geophysical modeling estimates low variance across the region.	Base-flow SC estimated by a geophysical model for the new area (195µS/cm) was 10.8% higher than modeled in the original area (176µS/cm). The model estimated that 76.2% of the new area has base-flow SC less than the background SC level in the original area (210 µS/cm). (++)
B.	Empirically estimated background	
5	Lowest observed SC in the new area is approximately as low as in the original area.	The minimum SC was 40 and 49µS/cm in the original and new areas, respectively. (+)
6	Similar SC 10^{th} or 25^{th} centiles of the full data sets in both areas.	The 10^{th} centile for the new area (163 µS/cm) is 26% higher than for the original area (120µS/cm). The 25^{th} centile for the new area (244 µS/cm) is 31% higher than for the original area (169µS/cm). (−)
7	The proportion of measured sites less than the background limit from the original area (210µS/cm) is not negligible.	From a mixed data set, 19% of unique sites have measured SC ≤210µS/cm. (+)
8	The median of reference sites in the new area is less than the background level for the original area.	NE. Confirmed reference sites not available from Ohio. Sites identified in the Ohio data sets as reference sites often had poor habitat scores. (0)
9	In mixed data sets, the lower quartile of the new area should overlap the lower quartile of the original area in most months.	Both the lowest quartiles and the interquartile ranges of SC values overlap during months when more than 30 unique sites were sampled. (+)
C.	Distribution of low SC sites	
10	Low SC sites (<210 µS/cm) are spatially dispersed.	Sites <210 µS/cm are widely distributed across Ohio Ecoregion 70, suggesting that they represent the regional background (Fig. 2c). (+)
11	Sites with SC less than a SC benchmark are spatially dispersed.	Sites ≤ 340 µS/cm are dispersed throughout the new area and represent 43.6% of unique sites in the new area (Fig. 2b). (+)
D.	Biological indicators of background SC	
12	Salt-intolerant genera (XC_{95}<340µS/cm) are found in the historical record.	Not examined. (0)
13	Salt-intolerant genera (XC_{95}<340µS/cm) are found in recent records.	Seven genera occur in the new area with XC_{95} values <340µS/cm. (++)
14	Regional biota is similar.	Of the 139 genera found in the original area, 90.6% are represented in the data set for the new area. (+)
E.	Relevance and reliability of evidence	
15	Larger data sets are more reliable.	The number of samples (4452) and sites (809) in the new area is sufficient to estimate background. (++)
16	Quality assured instream chemical measurements are more reliable.	Meter calibration and other good field and laboratory practices are required by Ohio EPA standard operating procedures. (+)
17	Appropriate reporting units are required.	Units are reported as SC. Other units and detection limits are as designated. (+)
18	Comparisons are appropriate only if ionic mixtures are similar.	Included samples in both areas are dominated by salts of HCO_3^- plus SO_4^{2-} (Cl^- represents less than 50% of anions by weight). (+)

	Types of evidence	Results
19	Comparisons are more reliable if collection methods and sampling windows are similar.	The new area data set was collected throughout the year but was much more temporally concentrated and was not spatially distributed by a probability design. (−)
20	Validation increases reliability.	Other data sets are available for the new area, but they are very small (<20samples), so no validation was attempted. (0)
21	Inclusion of likely background sites makes identification of background SC more probable.	Small catchments, which tend to be in higher elevation and less disturbed, are included in all data sets, but the proportion of such sites is small. The proportion of land cover in native vegetation is not known. (0)
22	Abrupt changes in SC values at political boundaries make it likely that differences are not natural.	There are no abrupt changes in SC measurements at the state borders that divide the new area from the rest of the ecoregion. (+)
23−27	Not applicable, because reference sites for the new area are not useable.	No results. (0)

SC = specific conductivity, NE = no evidence

Table 6

Weight of evidence (WoE) for each consideration and summary WoE for the hypothesis that the background specific conductivity (SC) is not different in Ohio.

Consideration	Summary WoE
Regional properties	The geology, climate, and physiography are similar throughout Ecoregion70. A geophysical model gave SC results that were similar for the two areas. Therefore, the geophysical processes that determine natural background are shared throughout the ecoregion. (+)
Empirically estimated background	The minima in the two areas are very similar, and 19% of sites in the new area are within the background level for the original area based on reference sites. However, the estimated natural background (25^{th} centile) is 31% higher in the new area than the original area. Because the sampling designs and timing are different, these results are ambiguous. (0)
Distribution of low SC sites	The spatial distribution of low SC sites is consistent with residual natural background for the area. However, the data set is not optimal in either coverage or sampling of smaller streams. (+)
Biological properties	Ninety percent of the genera in the original area occur in the new area. Salt intolerant genera are present; therefore, some places are believed to have low SC. However, species within some genera in the new area could have different salt tolerance from species in the rest of the ecoregion, so results are supportive but not conclusive. (+)
Relevance and reliability of evidence	The reliability in terms of data quantity and QA procedures of the Ohio Ecoregion 70 SC data set was judged to be very good, but relevance in terms of similarity of sampling is questionable. The most prominent differences in Ohio are the uneven spatial distribution of sampling, the high proportion of July–August sampling, and deployed substrate versus kick-net biological sampling. (0) However, both the biological data and the regional property information were judged to be highly relevant and reliable. (+ +)
Summary for the body of evidence	Based on SC sampling, background stream SC in the new area appears to be higher than in the original area, but the geophysical model indicates that the difference in natural background is small. Potential natural causes of the difference are not supported by regional properties and the biota similarities imply similar background SC. (+)