DE GRUYTER DOI 10.1515/jbca-2012-0005 — Journal of Benefit-Cost Analysis 2013; 4(1): 81–105

Patrick J. Walsh* and William J. Wheeler

# Water quality indices and benefit-cost analysis

**Abstract:** The water quality index (WQI) has emerged as a central way to convey water quality information to policy makers and the general public and is regularly used in US EPA regulatory impact analysis. It is a compound indicator that aggregates information from several water quality parameters. Several recent studies have criticized the aggregation function of the EPA WQI, arguing that it suffers from "eclipsing" and other problems. Although past papers have compared various aggregation functions in the WQI (usually looking at correlation), this is the first paper to examine these functions in the context of benefit-cost analysis. Using data from the 2003 EPA CAFO rule, the present paper examines four aggregation functions and their impact on estimated benefits. Results indicate that the aggregation method can have a profound effect on benefits, with total benefit estimates varying from $82 million to $504 million dollars. The net benefits of the rule vary from negative to positive over this range of estimates. Furthermore, a sensitivity analysis does not find convincing evidence to substitute the current aggregation function, although several changes to the underlying WQI methodology may be warranted.

**Keywords:** cost benefit analysis; valuation; water quality; water quality index.

***Corresponding author: Patrick J. Walsh,** US EPA, National Center for Environmental Economics, 1200 Pennsylvania Avenue, Ariel Rios Bldg, MC1809T DC Washington 20460, USA, e-mail: walsh.patrick@epa.gov
**William J. Wheeler:** US EPA, National Center for Environmental Economics, 1200 Pennsylvania Avenue, Ariel Rios bldg, MC1809T DC Washington 20460, USA

## 1 Introduction

There are several laws and regulations that require the quantification of water quality changes. For instance, the EPA is required to do benefit-cost analysis on economically significant rules, which requires quantification and monetization.[1] Additionally, Section 303(d) of the Clean Water Act requires states to report water

**1** See OMB's Circular A-4 for additional information: http://www.whitehouse.gov/omb/circulars_a004_a-4/.

quality conditions to the EPA. The quantification of water quality changes is inherently problematic; however, since there are a range of water quality indicators to choose from, which can vary in importance over geographic regions and represent different aspects of quality. Furthermore, it can be difficult to convey relevant water quality information to policy makers and the general public, who do not always have technical knowledge about the components of waterbody health.

To overcome these obstacles, analysts developed a Water Quality Index (WQI) (Brown, McClelland, Deininger and Tozer, 1970) to transmit complex water quality information. The EPA has used the WQI for the past few decades to quantify and monetize water quality changes in several of its Regulatory Impact Analyses (RIAs). It has also been used in multiple stated preference studies of water quality precisely because it easily communicates water quality information to a wide range of people.

The WQI is a composite indicator that combines information from multiple water quality parameters into a single overall value (on a 0–100 scale). This indicator has seen widespread use since its inception and is employed by multiple states and countries (Ott, 1978; Pesce and Wunderlin, 2000; Prakirake, Chaiprasert and Tripetchkul, 2009; Taner Üstün and Erdinçler, 2011). Creating the WQI involves three main steps (U.S. EPA, 2009): [1] obtain measurements on individual water quality indicators, [2] transform measurements into "subindex" values to represent them on a common scale, [3] aggregate the individual subindex values into an overall WQI value.

Given the widespread use of the WQI over three decades, it is surprising that more attention has not been paid to the construction of the index. A few sources have criticized EPA's current approach to step [3] (Dojlido et al., 1994; Cude, 2001), where the geometric mean is used to aggregate the subindex values, and others have proposed variations of the WQI (Landwehr and Deininger, 1976; Smith, 1990; Swamee and Tyagi, 2000; Gupta, Gupta and Patil, 2003). However, no past papers have examined the impact of WQI construction on benefit-cost analysis. The present paper demonstrates the effect of using four different WQI aggregation functions – the geometric, arithmetic, and harmonic means, and the minimum operator – on a benefit-cost analysis of a past EPA rule. Data from the EPA CAFO Rule (U.S. EPA, 2003a) is used to calculate national benefits under the four variations. Results indicate that the aggregation function can have a profound impact on the estimated benefits of the rule; yielding a range of $82 million to $504 million dollars. In fact, the choice of indicator determines whether the net monetized benefits of the CAFO Rule are negative or positive. Additionally, a sensitivity analysis supports the continued use of the geometric mean, while recognizing that several steps in the construction of the WQI need to be updated.

# 2 Background

In the latter half of the 20th century the United States initiated an expansive effort to protect and improve the nation's clean water, including the passage of the Federal Water Pollution Control Act (1948), the Clean Water Act (1972, with major amendments in 1977), and the Safe Drinking Water Act (1974). In order to justify the billions of dollars to be spent on water pollution programs and maintain public and private support, it quickly became clear that a method of communicating progress on program goals was needed (McClelland, 1974). Although several indices had been in use by state and regional entities already, the National Sanitation Foundation (NSF) was tasked with developing a tool that could be used more broadly, with the following objective: the method was to have a uniform, comprehensible format, which transparently conveyed progress in attaining program objectives. The main purpose of the tool was to communicate "water quality information to the lay public and to legislative decision makers." It was not intended to be used as a technical or scientific predictive model (McClelland, 1974). The NSF WQI was officially defined in Brown et al. (1970).

## 2.1 Use of the WQI in valuation

The WQI has since become one of the chief ways to communicate information about water quality, and is regularly used in the economic valuation literature. The WQI was regularly employed by early stated preference valuation studies because it can communicate complex water quality information to the general public. Pretesting and focus groups have indicated that survey participants can understand the WQI more easily than when individual pollutants are used (Carson and Mitchell, 1981; Eom and Larson, 2006).[2]

The WQI was first applied in water quality valuation studies by Mitchell and Carson (1989) and Smith and Desvousges (1986).[3] Both of these studies relied on Vaughan's (1981) transformation of the WQI into the water quality ladder (WQL) as an aid to respondents in their surveys. The WQL divides the 0–100 WQI range into

**2** While there is only limited evidence regarding how the general public perceives individual water quality measures, the evidence that does exist suggests that the linkage between objective water quality measures and perceived water quality is not always strong (see Binkley and Hanemann, 1978; Pendleton, Martin and Webster, 2001; Hoyer, Brown and Canfield Jr, 2004; Jeon, Herriges Kling and Downing, 2005). The adaptation of indexes to summarize water quality for the general public hopefully helps fill this gap.

**3** In stated preference surveys administered in 1980 and 1981, respectively.

five "designated use" classifications and was designed to better connect water quality information to policy goals.[4] Since these surveys, the WQI and the WQL have continued to be used in valuation surveys (Eom and Larson, 2006), while similar compound indices have been used in other stated preference studies (Johnston, Schultz, Segerson, Besedin and Ramachandran, 2012). The WQI has been used so often in stated preference studies that it has been applied to organize meta-analyses (Johnston, Besedin and Wardwell, 2003; van Houtven, Powers and Pattanayak, 2007) and benefit transfers (van Houtven, Pattanayak, Patil and Depro, 2011).

## 2.2 Indicators in the WQI

Over several iterations of the WQI since the 1970s, multiple water quality parameters have been used. Ideally, parameters which are both targeted by policy and representative of waterbody health would be used in the index. Under the Clean Water Act, states are required to biennially assess the attainment of designated uses in their waters and identify the sources of any impaired waterbodies.[5] Water quality problems from non-point sources, such as oxygen depletion, nutrients, and pathogens have consistently been cited in these reports as the primary causes of water quality impairment in the U.S. since the late 1970s.[6]

In light of the variety of state and national standards on water pollution, an assortment of water quality parameters have been used in past WQIs. Table 1 contains descriptions of the main water quality parameters commonly encountered, including dissolved oxygen (DO); biological oxygen demand (BOD); nutrients; sediment, silt and total suspended solids (TSS); pH; bacteria and pathogens; and temperature. The WQI is normally applied at a relatively fine spatial scale. For instance, McClelland (1974) calculated the WQI at individual monitoring stations. Recent EPA studies evaluate it at a similar spatial scale, normally the "reach" level (essentially an individual segment of a river or stream). Applying the WQI at this fine geographic level allows it to better account for spatial heterogeneity in quality.

**4** Designated uses are frequently used by states in water quality policy. The five designated uses in the WQL are: (1) acceptable for boating, (2) acceptable for rough fishing, (3) acceptable for game fishing, (4) acceptable for swimming, and (5) acceptable for drinking.

**5** The EPA website http://water.epa.gov/lawsregs/guidance/cwa/305b/index.cfm has these assessments since 1992. Earlier reports are available from EPA's National Service Center for Environmental Publications at http://www.epa.gov/nscep/index.html.

**6** A lack of a consistent reporting method in these reports prevents a quantitative presentation.

https://doi.org/10.1515/jbca-2012-0005 Published online by Cambridge University Press

**Table 1** Parameters used in past WQIs.

| Pollutant class | Measured as | Description |
|---|---|---|
| Low oxygen | Dissolved oxygen (DO), biological oxygen demand (BOD) | Oxygen in the water column is necessary for fish and other aquatic organisms and is generally measured in two ways; the first is dissolved oxygen (DO), a direct measurement. Some species can tolerate low DO but most desirable fish species need 3 or 4 mg/l while larvae and juveniles even more sensitive. Adults and juveniles can usually tolerate short time periods of low DO but prolonged episodes of low DO can create dead zones. The other measurement of oxygen is biological oxygen demand (BOD). Organic wastes (such as sludge and biosolids) are broken down in water by oxygen using-microorganisms; BOD is the amount of oxygen needed by these organisms. Water polluted by organic waste (or the creation of organic waste by toxics killing aquatic vegetation) will lead to growth in the population of the microorganisms as they feed, increasing BOD (and denying oxygen to other species). Low DO and high BOD can be caused by agricultural runoff (fertilizers and manure), municipal sewage, and other organic wastes |
| Temperature | Temperature | Not only do fish species require water in a certain temperature range to live, higher temperature can also increase BOD by stimulating microorganisms and fish respiration |
| Nutrients | Nitrogen (N), phosphorus (P) | Excess nutrients (nitrogen and phosphorus) can overstimulate the growth of aquatic vegetation and algae, which can clog waters, choke off submerged aquatic vegetation, and lead to oxygen depletion. Excess nutrients can come from agricultural and urban runoff (e.g., fertilizer from lawns); and municipal wastewater treatment plants |
| Sediment, Clarity | Total suspended solids (TSS), Clarity | Too much sediment (and silt) can suffocate eggs and larvae, interfere with recreational activities by reducing water clarity and filling in waterbodies, and may carry attached nutrients and toxic pollutants. Sediment and silt can be measured as total suspended solids, total dissolved solids, or their sum, total solids. Clarity can also be measured directly in turbidity units or via a Secchi disk, a standardized tool used to measure clarity depth. Sediments result from erosion of agricultural and urban nonpoint sources |
| Bacteria and pathogens | Fecal coliform, *E. coli* | Bacteria and pathogens such as fecal coliform and *E. coli* cause illnesses in humans (through direct contact or ingestion of water or shellfish). These pathogens come from municipal wastewater treatment plants, combined sewer overflows, urban stormwater, and livestock runoff. Because there are numerous pathogens, fecal coliform or *E. coli* are often used as an indicator of the presence of other pathogens |

(**Table 1** Continued)

| Pollutant class | Measured as | Description |
|---|---|---|
| pH | 0–14 Scale | The pH scale measures how acidic or basic (or alkaline) a water body is. A low score indicates acidity, a high score indicates basic, and a score of seven is neutral. Many biological processes cannot occur in excessively acidic or basic conditions. Problems with pH can be typically caused by mine drainage and tailings; and atmospheric deposition |

## 2.3 WQI construction

There are currently two main types of WQI: relative and absolute. Relative indices focus on the achievement of legislated thresholds or criteria. For example, Carruthers and Wazniak (2004) formulate a WQI based on the achievement of ecosystem criteria. Several binary variables, indicating criteria achieved or not, are averaged together to form their WQI. Absolute indices, on the other hand, are independent of criteria or thresholds, and wholly based on water quality measurements. The present paper focuses on absolute indices, since they can be applied more broadly than relative indices and are more commonly used by the EPA and other environmental agencies.

Ott (1978) reviewed the existing indices and literature and formulated a list of 20 criteria for an ideal WQI, which are contained in Table 2. Although no index is expected to meet all 20 (particularly since some of them could be contradictory), they provide a set of dimensions along which WQIs can be evaluated. The criteria are relatively expansive, and different weight should be given to different criteria. Consequently, they do not provide an objective method of evaluation, but they at least provide a model to aspire towards, where a wide number of violations indicate distance from the ideal.

One of the earliest WQIs appeared in Horton (1965), which developed a compound index of 10 water quality variables. This was the first paper to outline the three main steps for WQI construction discussed earlier: [1] select quality characteristics and obtain measurements, [2] establish a rating scale for each characteristic and transform observations into subindex values, and [3] select a weighting method and aggregate individual subindex values into one index number. Horton's approach to steps [1] and [2] was rather arbitrary, so the index was not useful for policy analysis. For step [3], Horton used the arithmetic aggregation

**Table 2** Ott (1978) criteria.

| No. | Criteria |
|---|---|
| 1. | Relatively easy to apply |
| 2. | Strikes a reasonable balance between oversimplification and technical complexity |
| 3. | Imparts an understanding of the significance of the data it represents |
| 4. | Includes variables that are widely and routinely measured |
| 5. | Includes variables that have clear effects on aquatic life, recreational use, or both |
| 6. | Includes toxic substances |
| 7. | Can easily accommodate new variables |
| 8. | Based on recommended limits and water quality standards |
| 9. | Developed from a logical scientific rationale or procedure |
| 10. | Tested in a number of geographical areas |
| 11. | Shows reasonable agreement with expert opinion |
| 12. | Shows reasonable agreement with biological measures of WQ |
| 13. | Dimensionless |
| 14. | Has a clearly defined range |
| 15. | Exhibits desirable statistical properties permitting probabilistic interpretations to be made |
| 16. | Avoids eclipsing |
| 17. | Shows sensitivity to small changes in WQ |
| 18. | Applicable for showing trends over time, for comparisons of different locations, and for public information purposes |
| 19. | Includes guidance on how to handle missing values |
| 20. | Limitations of the index are clearly documented |

function (or mean), but with temperature and "obvious pollution" entering multiplicatively, as seen in equation (1)[7]:

$$WQI_{Horton}=WQI_A{*}T{*}O \tag{1}$$

where temperature is $T$, obvious pollution is $O$, and $WQI_A$ is the arithmetic weighting of the other water quality variables, illustrated in equation (2):

$$WQI_A=\sum_{i=1}^{n} q_i w_i \tag{2}$$

In (2), $q_i$ is the 0–100 rating for each variable, $n$ is the total number of water quality parameters, and $w_i$ are the weights, where $\sum_{i=1}^{n} w_i=1$.

7 "Obvious pollution" indicates areas with conditions that are offensive to sight and smell, such as oil slicks, debris, and scum and sludge deposits. Horton had obvious pollution and temperature enter multiplicatively because they "cannot readily be rated to show gradations in quality but fall more into the category of 'yes' or 'no' indicators." However, obvious pollution is somewhat subjective (and temporally dependent), and was not included in later WQIs.

**Table 3** Professional fields of invited panelists for NSF WQI.

| | |
|---|---|
| Regulatory officials (federal, interstate, state, territorial and regional) | 101 |
| Local public utilities managers | 5 |
| Consulting engineers | 6 |
| Academicians | 26 |
| Others (industrial waste control engineers and representatives of professional organizations) | 4 |
| Total | 142 |

Source: (Brown et al., 1970).

Following the work of Horton, the NSF created the seminal WQI several years later, as published in Brown et al. (1970). To accomplish steps [1] and [2] of WQI construction in a less arbitrary way than Horton, a Delphi survey of 142 water quality experts was performed, with the composition of this group appearing in Table 3.[8] The experts were asked to evaluate the importance of a wide variety of water quality indicators, with chances to re-evaluate their scores in additional rounds of mailings. Once the expansive list of indicators was reduced to nine, respondents were asked to create graphs that translated variable concentrations into 0–100 values, with higher values indicating healthier water. The final versions of these graphs were used to produce the subindex values used in step [3] of the WQI construction – the focus of the present paper. An example of the dissolved oxygen (DO) subindex curve from Brown et al. (1970) appears in Figure 1. This graph shows the non-linear relationship between DO saturation and water quality (the subindex values were originally called "Q-values", which appear on the y-axis). The subindex curves from Brown et al. (1970) are still used in current EPA analysis, as well as in other studies that use the WQI (Johnston et al., 2005). In the third round of the Delphi survey, after the field of most preferred parameters was reduced to nine, respondents were asked to rate the parameters on a scale of 1–5. This information was used to create the subindex weights used in the aggregation functions. The subindex weights in recent EPA RIA's were derived directly from the nine-parameter weights in Brown et al. (1970).

**8** The Delphi method is a structured interview of experts used to quantify uncertainty, which was developed by the Rand Corporation (Dalkey, 1968; Morgan and Henrion, 1990). In the three-round Brown et al. (1970) Delphi survey, the first round presented 35 water quality indicators to respondents for evaluation, and gave them a chance to recommend other indicators. The second round included these new indicators, as well as the evaluation results from the first round, and asked participants to indicate their opinion of the "most important" indicators. The third round involved rating the importance of each indicator.

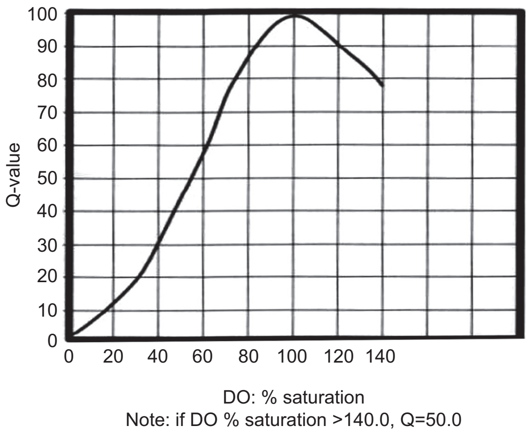


**Figure 1** Subindex curve for DO.
Source: (U.S. EPA, 2003b).

Similar to Horton (1965), Brown et al. (1970) also employed basic arithmetic weighting ($WQI_A$), although without the multiplicative *T* and *O* variables. The arithmetic weighting is the most transparent, and is still used today in several countries as the official method, including Turkey (Taner et al., 2011) and Argentina (Pesce and Wunderlin, 2000). In fact, the full NSF WQI developed in Brown et al. (1970) was adopted by the Thai Ministry of Natural Resources and Environment in 1995 as the main tool for water assessment (Prakirake et al., 2009).

Several years after Brown et al.'s (1970) study, McClelland (1974) introduced a different form of weighting to the WQI: the geometric mean. McClelland was concerned that the arithmetic mean lacked sensitivity to low value parameters, a characteristic later deemed "eclipsing" (16 in Ott's list). Eclipsing can cause problems with the original WQI objective of "conveying progress in program objectives." For instance, if a water quality program successfully targets a low value parameter and the WQI is not sensitive to the resulting improvement, policy makers may erroneously shift funds and efforts away from that program. McClelland instead proposed the weighted geometric mean appearing in (3), where $q_i$ and $w_i$ are as defined before:

$$WQI_G = \prod_{i=1}^{n} q_i^{w_i} \tag{3}$$

To compare the arithmetic WQI to the geometric WQI, McClelland obtained survey responses from over 100 water quality experts – with 30 of them having participated in the original Brown et al. (1970) survey. The experts were given data from actual stream samples and asked to rate them in three iterations similar

https://doi.org/10.1515/jbca-2012-0005 Published online by Cambridge University Press

to a Delphi procedure. When compared to the experts' ratings of waterbodies, the arithmetic WQI averaged 10–15 units higher than the experts' evaluation, whereas the geometric WQI averaged only 6 units different, distributed above and below the experts. Using a similar process, Landwehr and Deininger (1976) also found that the geometric mean matched experts' ratings better than several other WQI variations. Although it violates a few of the lesser criteria from Ott (6, and potentially 15, 17, and 19), it satisfies the majority of them. The geometric mean has been used in all EPA RIA's that employ a WQI, including: the CAFO rule (U.S. EPA, 2003a), the Concentrated Aquatic Animal Production rule (U.S. EPA, 2004a), the Meat and Poultry Processing Rule (U.S. EPA, 2004b), the Construction and Development Rule (U.S. EPA, 2009), and the Florida Numeric Nutrient Criteria Rule (U.S. EPA, 2010).

The third WQI aggregation method explored in this paper uses the square root of the harmonic mean of squares, and was first popularized in Dojlido et al. (1994). This mean, appearing in equation (4), does not use weights for the individual indicators.[9]

$$WQI_H = \sqrt{\frac{n}{\sum_{i=1}^{n} 1/q_i^2}} \tag{4}$$

Dojlido et al. (1994) found that this specification was more sensitive to the most impaired indicator than the arithmetic or geometric means – reducing eclipsing – while still accounting for the influence of other indicators. One concern with (4), however, is that it may result in "ambiguity," a situation where all subindices indicate good water quality but the overall indicator does not (Swamee and Tyagi, 2000). So although the harmonic variation makes potential improvements in Ott's 16th criteria, it may fall short on numbers 8–12, and 17–19. The harmonic mean was also recommended by Cude (2001), who developed a WQI for the State of Oregon. Drawing from Dunnette (1979), Cude (2001) also popularized ecoregion-specific subindex curves, allowing the WQI to be tailored to local conditions. This regional approach to the subindex curves was used in the analyses for EPA's Construction and Development Rule (U.S. EPA, 2009) and the Florida Numeric Nutrients Rule (U.S. EPA, 2010). However, the harmonic mean itself has not been used in EPA regulatory analysis.

The final subindex aggregation method is the minimum operator, which has been proposed as another method to eliminate eclipsing. As shown in equation (5), the overall WQI in this variation is simply the lowest subindex value:

$$WQI_M = \min(q_1, q_2, \dots q_n) \tag{5}$$

9 $WQI_H = 0$ if $q_i = 0$ for any $i$.

https://doi.org/10.1515/jbca-2012-0005 Published online by Cambridge University Press

Smith (1990) established widespread interest in this method; arguing that the limiting indicator is critical information hidden by other aggregation methods. It has been popular with government environmental organizations in New Zealand (Nagels, Davies-Colley and Smith, 2001) and Canada (Khan, Husain and Lumb, 2003). Nagels et al. (2001) argue that it is particularly important for certain designated uses, like primary contact recreation. However, the minimum operator is totally insensitive to changes in the other variables, and so is not useful for monitoring purposes or for comparing two waterbodies (Swamee and Tyagi, 2000). Returning again to the WQI objective of "conveying progress in obtaining program objectives", the $\mathrm{WQI}_{\mathrm{M}}$ approach fails unless the program solely targets the minimum indicator (and, similar to the harmonic variation, may not fully satisfy criteria 8–12 and 17–19 from Table 2). Other papers that use or support the minimum operator include Pesce and Wunderlin (2000), Flores (2002), Parparov, Hambright, Hakanson and Ostapenia (2006), Simões, Moreira, Bisinoti, Gimenez and Yabe (2008), and Prakirake et al. (2009).

If a normal harmonic mean had been used – as opposed to the square root of the harmonic mean of squares – all of these aggregation functions could be derived from the generalized mean (also called the power mean or Hölder mean), appearing in (6):

$$M_p = \left( \sum_{i=1}^{n} w_i q_i^{\,p} \right)^{1/p} \tag{6}$$

When $p$ takes the value of 1, the mean is arithmetic, when $p$=0 it is the geometric mean (using L'Hopital's Rule as $p$ approaches 0), when $p$=–1 the normal harmonic mean appears, and finally when $p$=–∞ it is the minimum operator.[10] The arithmetic ($\mathrm{M}_A$), geometric ($\mathrm{M}_G$), and harmonic ($\mathrm{M}_H$) means also have the following property,

$$\mathrm{M}_A \geq \mathrm{M}_G \geq \mathrm{M}_H, \tag{7}$$

However, the harmonic variation from Dojlido et al. (1994) does not fit into this inequality (Mercer, 2003), leaving us with $\mathrm{WQI}_{\mathrm{A}} \geq \mathrm{WQI}_{\mathrm{G}}$.

Other, more exotic, aggregation methods have also been proposed. For instance, Kung, Ying and Liu (1992) and Chang, Chen and Ning (2001) support the use of "fuzzy" evaluation tools to account for uncertainty in data sampling (from things like uneven weather patterns or infrequent sampling) and complexity in decision-making. Chang et al. (2001) evaluate their approach on a

**10** The generalized mean has been used extensively in economics, in particular in the field of price indices. Diewert (1993) provides background on how various means have been used in economic applications, with a focus on symmetric means.

Taiwanese river system and find that the traditional WQI provides a more pessimistic vision of water quality. However, the study does not provide sufficient evidence in favor of the fuzzy techniques over the traditional WQI.

Walski and Parker (1974) and Bhargava (1983) use “sensitivity functions” along with parameter weights to aggregate variables in WQI variations. These approaches are motivated by a desire to better connect the WQI value to the designated use of the water. However, Walski and Parker’s approach only considered recreational water quality, and their sensitivity functions had difficulty reflecting the proper weightings for each variable (Bhargava, 1983). Bhargava’s approach was designed to assess water quality and assign designated uses to different parts of a waterbody. The development of the WQL by Vaughan (1981) is favored in the literature for the latter purpose, without adding the complexity of the sensitivity functions. Overall, these more exotic types of indicators have not gained as much traction in the literature or in applied policy settings as the four WQI variations discussed here. While it is not completely clear why these and other (e.g., Sarkar and Abasi, 2006) approaches to the WQI have not been widely adopted, they add considerable complexity without significant demonstrated improvement. They therefore stray from the original goals of the WQI, since their construction is not easily communicated to policy makers or the general public.

The WQI was first applied by the EPA in benefit-cost analysis in U.S. EPA (2000), in a study examining the value of reductions in conventional pollutants (BOD, TSS, DO, and fecal coliform) arising from the Clean Water Act. It has since been used by EPA in several rules, as recently as 2010 (U.S. EPA, 2010).[11] Originally, EPA applied the Mitchell and Carson (1989) (or the later version, Carson and Mitchell, 1993) values directly to modeled changes in WQI, but now relies on meta-analyses of valuation studies (U.S. EPA, 2009). The next two sections present a more thorough examination of past EPA approaches.

# 3 Data

This paper uses water quality data from the EPA RIA for the 2003 CAFO rule (U.S. EPA, 2003a). The rule uses National Pollutant Discharge Elimination System (NPDES) permits, effluent limitations, and technology standards to protect water quality from manure, wastewater, and other process waters generated by CAFOs. The data contain baseline and projected measures of six water quality variables: biological oxygen demand (BOD), dissolved oxygen (DO), fecal coliform bacteria

**11** See Griffiths et al. (2012) for more background on the approaches to benefits estimation in EPA water rules.

https://doi.org/10.1515/jbca-2012-0005 Published online by Cambridge University Press

(FEC), total suspended solids (TSS), nitrogen ($NO_3$), and phosphorous ($PO_4$).[12,13] These six variables were used by U.S. EPA (2003a) to create a six parameter WQI, with weights rescaled from the nine-variable McClelland (1974) WQI.[14] The water quality data are geocoded at the RF3 Lite network level, and include 1,817,988 reaches totaling 2,655,437 miles within the contiguous 48 states.[15] The projected values were obtained through several water quality models, based on local conditions and projected impacts of the policy.[16]

# 4 Analysis

This exercise is aimed at measuring the sensitivity of estimated benefits to the specification of the aggregation function. First, the baseline and projected water quality parameters are transformed into subindex values. Next, the subindex values are fed to the four different aggregation functions to calculate WQIs. The weights for the geometric and arithmetic functions are directly from U.S. EPA (2003a), and appear in Table 4. Figure 2 illustrates the distribution of national WQI scores for baseline water quality before the CAFO rule, for each WQI variation. The graphs exhibit considerably different pictures of water quality, illustrating the importance of the aggregation function to the WQI. For instance, the distributions of the harmonic mean and minimum index portray a much bleaker state than the arithmetic and geometric WQI variations. The distributions of the geometric and arithmetic WQIs illustrate the result from equation (7), with the arithmetic distributed higher than the geometric. The forecast of the change in

**12** Note that these are the baseline and projected values at the time of the rule, so have not been updated or changed since the rulemaking.

**13** We follow the terminology of the CAFO rule and refer to “baseline” water quality as the water quality before the policy, and refer to the water quality after the policy as the “projected” water quality.

**14** Three variables from McClelland’s analysis were therefore omitted: pH, temperature, and total solids. The weights are rescaled so that the ratios of the weights are retained and the weights still sum to one.

**15** RF3 lite, or Reach File 3 lite, is a subset of the Reach File 3 hydrologic database. The Reach File databases contain data on US surface waters, and are inputs to several large scale hydrologic models. The RF3 lite subset contains streams longer than 10 miles, as well as the small streams needed to connect those (>10 mile) segments. For additional information, see U.S. EPA (2003b).

**16** For additional information about the 2003 CAFO rule, see http://www.epa.gov/npdes/regulations/cafo_fedrgstr.pdf and for the water quality benefits estimation in particular, http://cfpub.epa.gov/npdes/docs.cfm?view=allprog&program_id=7&sort=name#cafofinalrule_nationaleconbenefits_2003.

**Table 4** WQI weights.

| Parameter | Weight |
|---|---|
| BOD | 0.15 |
| DO | 0.24 |
| FEC | 0.23 |
| TSS | 0.1 |
| $NO_3$ | 0.14 |
| $PO_4$ | 0.14 |

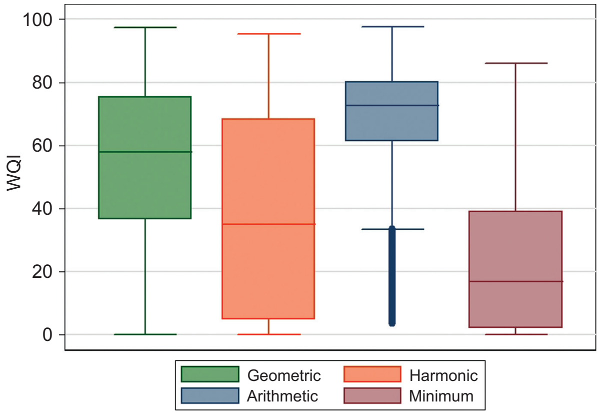


**Figure 2** National WQI distributions in CAFO baseline.

water quality is also heavily influenced by the aggregation function. Table 5 contains summary statistics on the change in WQI from baseline to projected for each WQI variation. The harmonic mean, which is designed to be most sensitive to the lowest parameter, shows the greatest change, while the arithmetic WQI shows a smaller change that is concentrated around zero. The different response by these two functions provides an example of eclipsing.

Another important policy consideration is the ranking of waterbodies. While the absolute level of the water quality index is important, state planning typically involves targeting the lowest performing waterbodies for cleanup. If the ranking from the different indices varies substantially, they could result in significantly different cleanup actions. At the national level, the CAFO baseline ranking changes considerably between indices. When the waterbodies are ranked from best to worst, the average absolute change in ranking from the geometric index to the arithmetic index is 33,375, to the harmonic index is 17,039, and to the minimum index is 37,023. However, due to the large number of waterbodies

https://doi.org/10.1515/jbca-2012-0005 Published online by Cambridge University Press

**Table 5** Policy forecast for change in WQI.

| WQI | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|
| Geometric | 0.2983 | 1.4793 | −46.4499 | 52.0442 |
| Harmonic | 0.4577 | 2.5227 | −75.0607 | 77.2153 |
| Arithmetic | 0.1004 | 0.5325 | −25.8412 | 23.2690 |
| Minimum | 0.3072 | 1.9282 | −69.1100 | 65.9200 |

(n=577,068).

nationally, a small index change can result in a large ranking change, so it may be best to compare rankings at the state level. In Ohio, for example, which has 11,838 reaches, the average difference in waterbody ranking between the geometric index and the arithmetic, harmonic, and minimum indices is 797, 322 and 884, respectively.[17] So even at the state level there is still considerable movement. If the ordinal ranking of waterbodies is a determinant of pollution control budgets, the choice of index could affect the distribution of funds and programs.

The next step in the analysis is the monetization of the WQI changes. The CAFO rule was projected to yield "moderate" changes in the six WQI variables. Table 6 contains the changes in loadings to the agricultural cells in the water quality model used in the CAFO analysis (U.S. EPA, 2003a).[18] Following the CAFO RIA, the projected change in water quality is valued for each state using the following benefit transfer function estimated by Carson and Mitchell (1993) in their national contingent valuation study of water quality, appearing in equation (8).[19,20]

$$\begin{aligned}\Delta TOTWTP=&\exp[0.8341+0.819*\log(WQI_1/10)+0.959*\log(Y)]\\&-\exp[0.8341+0.819*\log(WQI_0/10)+0.959*\log(Y)]\end{aligned} \tag{8}$$

17 We thank an anonymous reviewer for pointing out this important consideration.

18 These represent the total national CAFO loadings actually distributed to agricultural cells and production area loads input directly into the reaches. The model uses these loads to estimate the various water quality parameters in each area. DO was not directly reported in the CAFO documentation since the water quality model derives it from the other parameters. See U.S. EPA (2003b) for more detail.

19 As in the CAFO RIA, the WQI for each state is calculated by weighting each reach by its length as a proportion of the total reach miles in the state. Once the statewide change in water quality is calculated, that value is plugged into the benefit transfer function in equation (7).

20 Carson and Mitchell (1993) used a national, in-person stated preference survey to ask respondents to value changes in the WQI anchored to achievement of the goals of the Clean Water Act (that is, fishable and swimmable water). The focus of the survey was a national change in water quality, similar to the CAFO rule.

**Table 6** Estimated loadings and reductions to RF3 lite network, CAFO rule.

| | Nitrogen (lbs/yr) | Phosphorus (lbs/yr) | Sediment (lbs/yr) | BOD (lbs/yr) | Fecal coliforms (MPN/yr) |
|---|---|---|---|---|---|
| Baseline loads | 165.7 million | 243.5 million | 47.54 billion | 60.93 million | 6.46 E+23 |
| Post-regulatory loads | 149.4 million | 209.1 million | 46.61 billion | 46.10 million | 5.676 E+21 |
| Removals | 16.27 million | 34.41 million | 933.4 million | 14.74 million | 7.8 E+20 |

Source: U.S. EPA (2003a).
Note: Numbers may not add due to rounding.

**Table 7** National benefits, CAFO rule.

| WQI | Benefits |
|---|---|
| Geometric | $287,400,162.45 |
| Harmonic | $503,741,769.09 |
| Arithmetic | $81,913,882.61 |
| Minimum | $358,633,241.42 |

where $\Delta TOTWTP$ is the change it total household willingness-to-pay for a change in water quality, $WQI_1$ is projected WQI, $WQI_0$ is baseline WQI, and $Y$ is statewide annual household income.[21] The state values are aggregated to obtain the national estimate of benefits.

Table 7 contains the total estimated benefits of the CAFO rule for each WQI variation. These benefits vary from a low of $82 million for the arithmetic WQI to a high of $504 million for the harmonic WQI, exhibiting a six-fold difference.[22] EPA's favored geometric mean actually has the second-smallest estimated benefits, although the ordinal relationships in magnitudes might not hold in different places on the marginal benefit curve. The aggregation function therefore has a

**21** The approach follows the CAFO analysis, with figures inflated to 2001 dollars using the CPI. Note also that the published version of (8) includes covariates for household use and the importance of controlling pollution, as expressed by respondents. In the CAFO analysis, EPA used the Carson and Mitchell sample averages as a scalar value for the entire sample and incorporated the scalar value times the coefficient (for each variable) into the constant term.

**22** Note that the monetized benefit figures in Table 7 are not proportional to the mean WQI changes from the previous Table. This is due to the nonlinearity of the TOTWTP function appearing in (8).

surprisingly large impact on estimated benefits, which could have a large impact on policy recommendations derived from the benefit-cost analysis. Although previous papers have shown differences in the calculated WQI values as a result of the aggregation function, this paper is the first to estimate the impact on estimated benefits.

## 4.1 Sensitivity exercise

To further investigate the four aggregation methods, a hypothetical sensitivity analysis is performed on the Ohio water quality data from the CAFO Rule. Ohio is used because the state has relatively good monitoring data covering the majority of areas in the state.[23] Concentrating on one state should isolate subindex aggregation issues from other concerns with population, income, and heterogeneity in water quality monitoring.

A hypothetical water quality improvement is instituted in all waterbodies in the state to gauge the impact on estimated benefits for the four versions of WQI. Two different changes are analyzed for each water quality variable: a 5% increase and a five point increase in the subindex value.[24] The change in WQI is then monetized using the benefit transfer function from equation (8).

Figure 3 shows the change in benefits from increasing each variable by five points. The graph is dominated by the improvement in fecal coliform, valued at a maximum of over $160 million using the harmonic WQI. The substantial benefits of fecal coliform occur because it is the most impaired indicator in the majority of waterbodies, with an average subindex value of around two (out of 100). A five point jump is therefore a comparatively large improvement. Since the harmonic mean was designed to better account for the lowest value indicator, it is expected to place a higher value on this change. The eclipsing problem can be seen in this Figure with the arithmetic function. While the geometric, harmonic, and minimum WQIs experience a large jump in Figure 3, the arithmetic response is much more muted (or, eclipsed). Note that the geometric WQI does not appear to

23 Furthermore, the state has a diverse set of waterbodies. "Ohio is a water-rich state with more than 25,000 miles of streams and rivers, a 451 mile border on the Ohio River, more than 5,000 lakes, ponds, and reservoirs (>1 acre), and 236 miles of Lake Erie shoreline. Ohio has 10 scenic rivers comprising more than 629 river miles, the fourth largest total of any state in the nation," from http://www.epa.ohio.gov/dsw/general.aspx.

24 Each variable is increased individually, not compounded on top of the changes in other variables. Other changes in magnitude were also analyzed. However, the results were qualitatively similar to the 5% and 5 point changes, so are not presented.

https://doi.org/10.1515/jbca-2012-0005 Published online by Cambridge University Press

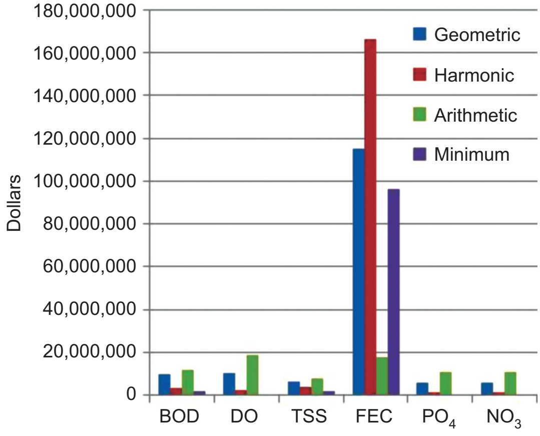


**Figure 3** Five point increase in each variable.

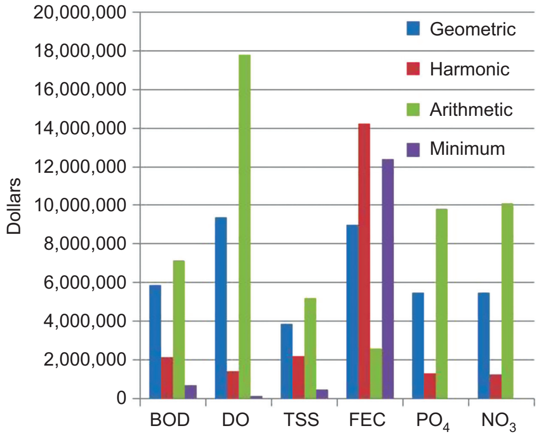


**Figure 4** Five percent increase in each variable.

be susceptible to eclipsing in this Figure, as it is relatively responsive to changes in FEC.

The more realistic five percent changes appear in Figure 4. The harmonic and minimum WQI are still quite sensitive to the improvement in FEC, and the arithmetic WQI still exhibits eclipsing. However, the minimum and harmonic WQIs are not particularly responsive to variables other than FEC. For example, since there were no reaches with nitrogen and phosphorous as the most impaired vari-

https://doi.org/10.1515/jbca-2012-0005 Published online by Cambridge University Press

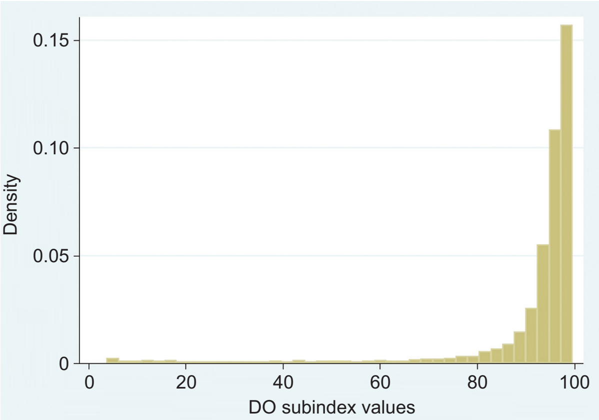


**Figure 5** Ohio DO subindex value distribution.

able, the minimum WQI assigns their improvements a value of zero. This is an unattractive quality for benefit cost analysis, since considerable efforts to reduce nutrient pollution would not be represented by corresponding increases in the WQI.[25]

Not only does the arithmetic mean WQI eclipse the change in FEC, it places a very high value on the change in DO. Figure 5 contains a graph of the distribution of DO subindex values in Ohio. The DO variable has the most right-skewed distribution of all indicators, with most reaches containing subindex values above 90. This highlights a particularly undesirable property of the arithmetic function: the level of the subindex value matters. This issue is further magnified for DO, since its subindex value is assigned the highest weight (Table 4), at 0.24.

The geometric WQI does not appear to suffer from eclipsing, and does not depend on the numerical level of the subindex value. Also, the ordering of the benefit values in the Figure aligns exactly with the parameter weights from Table 4. DO has the highest weight (0.24), so increases in it produce the greatest benefits – although only slightly higher than FEC (with a weight of 0.23). Improvements in individual parameters therefore yield benefits that are proportional to the weights assigned by the original panel of experts, following the important characteristic 11 from Ott's (1978) list.

**25** Particularly considering the millions of dollars currently being spent on combating nutrient pollution. For example, in Florida (U.S. EPA, 2010) and in the Chesapeake Bay (http://www.epa.gov/chesapeakebaytmdl/).

# 5 Conclusion

The WQI has become a central part of many EPA RIAs and is also widely used by environmental agencies in other countries (Nagels et al., 2001; Khan et al., 2003; Liou, Lo and Wang, 2004). However, the specification of the WQI and its impact on policy analysis has previously received scant attention. The present paper analyzes an important step in WQI construction, where several water quality variables are standardized and then aggregated into an overall value. Several recent studies (Dojlido et al., 1994; Cude, 2001) have criticized the use of the geometric mean – which is currently used in EPA RIAs – as the aggregation function. Although other papers have previously compared the statistical properties of different aggregation functions, this is the first paper to analyze the problem in the context of benefit-cost analysis.

Four aggregation functions were analyzed, which were selected because they have been used or supported in regulatory analysis for national or state entities. They include the arithmetic, geometric, and harmonic means, as well as the minimum operator. Data from the EPA CAFO Rule (U.S. EPA, 2003a) was used to estimate the benefits of a proposed water quality change for all four WQI variations. Additionally, two hypothetical changes in water quality were instituted to further examine the behavior of the four aggregation functions.

From the CAFO data, it is clear that estimated benefits are quite sensitive to the subindex aggregation function. Over the four different functions, benefits range from \$82 million to \$504 million. The geometric mean, which is used in EPA RIAs, sits near the middle of that range at \$287 million. Although these monetized benefits need to be added to other monetized benefits, such as the value of reduced nitrification of private wells (\$30.9–\$45.7 million), reduced public water treatment costs (\$1.1–\$1.7 million), and reduced livestock mortality (\$5.3 million), they represent the lion's share of monetized benefits. Since the total social costs of the rule were estimated to be \$335 million (U.S. EPA, 2003a), the choice of the aggregation function could move the rule from positive monetized net benefits to negative. Policy recommendations from the benefit-cost analysis could therefore vary drastically depending on the aggregation function used.

The sensitivity analysis did not support a switch from the geometric mean. With the geometric WQI, the importance of individual parameters to estimated benefits is a good reflection of the weights provided by a panel of hydrology experts. Also, the geometric mean does not inflate high valued indicators or eclipse the most impaired indicator as much as the other aggregation functions. The harmonic and minimum functions were found to be extremely sensitive to the most impaired variable, at the cost of others, while the arithmetic mean was subject to eclipsing and is dependent on the numeric level of water quality.

One of the main goals of the WQI is to communicate progress in achieving water quality program objectives, providing a signal to policy makers and the public. The aforementioned problems of eclipsing and sensitivity to the most impaired variable could provide the wrong signal and result in a misallocation of resources. On the one hand, a WQI that suffers from eclipsing would direct funds and project efforts away from the lowest valued indicator, since its improvement would not likely be reflected in the WQI. On the other hand, the harmonic and minimum indicators could potentially divert disproportionate funds to the lowest valued indicator, since a small improvement in its value would result in large increases in the WQI. If there is a strong desire to appropriately consider the low valued parameters (such as in Oregon where the harmonic indicator is favored), a more balanced approach would use the geometric WQI but also report the value of the most impaired indicator.[26]

Although this paper does not support a move from the geometric mean, the results highlight two pressing issues in the field. First, an updating of the WQI weights and subindex curves may be in order. The weights had a relatively strong influence on estimated benefits in the sensitivity analysis. Since the weights presently used by the EPA are based on a survey from the 1970s (Brown et al., 1970); an update may be needed. The biology, ecology, and limnology underlying water quality analysis have all improved in the last 40 years and expert opinion has likely evolved as well. Furthermore, most state and national water quality criteria have become more refined to different uses and there are now additional criteria for different pollutants.[27] Some of the criticism of the current WQI may be assuaged by developing new weights and subindex curves. Because WQI weights and subindex curves can be expected to be spatially-variant due to local conditions and preferences, a regional approach to the subindex curves, popularized by Cude (2001), represents a promising future path. That approach has already been used in the Construction and Development (2009) and Florida Numeric Nutrients (2010) rules, and has been met with widespread approval.

The second issue deals with comparisons across different WQI formulations. Although the literature and RIA's treat a unit change in WQI as the same, no matter the aggregation function, these variations may be measuring different things. The minimum aggregation function, for instance, is most concerned with the most degraded indicator. Individuals may value an improvement in the worst indicator in a stream differently than they value a change in the average of values across

26 Thanks to an anonymous reviewer for pointing out this approach.

27 It may also be desirable to convene a more diverse set of experts for a Delphi survey, since the previous panel represented in Table 3 is heavily influenced by regulatory officials.

indicators. In the field of hedonic analysis, Michael, Boyle and Bouchard (2000) find that different representations of water clarity – such as summer minimum, ten year average minimum, or maximum – yield different implicit prices. Future research should explore whether the different versions of the WQI are valued differently by survey respondents. Comparisons across the WQI variations may be affected by this issue, and a better understanding could improve future discussions about the WQI.

**Author note:** The views expressed in this paper are those of the author(s) and do not necessarily represent those of the U.S. Environmental Protection Agency. Thanks to comments from Charles Griffiths, Matt Massey, Julie Hewitt, and Todd Doley, as well as participants at the 2010 NAREA Annual conference and the 2010 Benefit-Cost Analysis Conference.

## References

Bhargava, D. S. (1983). Use of water quality index for river classification and Zoning of Ganga river. *Environmental Pollution Series B, Chemical and Physical, 6*(1), 51–67.

Binkley, C. S., & Hanemann, W. M. (1978). *The recreation benefits of water quality improvement: analysis of day trips in an urban setting*. Report No. EPA-600/5-78-010. Washington, DC: U.S. Environmental Protection Agency.

Brown, R. M., McClelland, N. I., Deininger, R. A., & Tozer, R. G. (1970). A water quality index – do we dare? *Water and Sewage Works, 11*, 339–343.

Carruthers, T., & Wazniak, C. (2004). Development of a water quality index for the Maryland Coastal Bays. In: C. E. Wazniak & M. R. Hall (Eds.), *Maryland's Coastal Bays: Ecosystem Health Assessment 2004*. DNR-12-1202-0009. Annapolis, MD: Maryland Department of Natural Resources Tidewater Ecosystem Assessment.

Carson, R. T., & Mitchell, R. C. (1993). The value of clean water: the public's willingness to pay for boatable, fishable, and swimmable quality water. *Water Resources Research, 29*(7), 2445–2454.

Chang, N. -B., Chen, H. W., & Ning, S. K. (2001). Identification of river water quality using the fuzzy synthetic evaluation approach. *Journal of Environmental Management, 63*(3), 293–305.

Cude, C. G. (2001). Oregon water quality index: a tool for evaluating water quality management effectiveness. *JAWRA Journal of the American Water Resources Association, 37*(1), 125–137.

Dalkey, N. C. (1968). *The Delphi method: an experimental survey of group opinion*. RAND Corporation Paper Series RM-5888-PR. Available at: http://www.rand.org/pubs/research_memoranda/RM5888.html.

Diewert, W. E. (1993). Symmetric means and choice under uncertainty. *Essays in Index Number Theory* (Vol. 1), North Holland, London.

Dojlido, J., Raniszewski, J., & Woyciechowska, J. (1994). Water quality index applied to rivers in the vistula river basin in Poland. *Environmental Monitoring and Assessment, 33*(1), 33–42.

Dunnette, D. A. (1979). A geographically variable water quality index used in Oregon. *Journal of the Water Pollution Control Federation, 51*(1), 53–61.

Eom, Y. S., & Larson, D. M. (2006). Improving environmental valuation estimates through consistent use of revealed and stated preference information. *Journal of Environmental Economics and Management, 52*, 501–516.

Flores, J. C. (2002). Comments to the use of water quality indices to verify the impact of Cordoba City (Argentina) on Suquia river. *Water Research, 36*(18), 4664–4666.

Griffiths, C., Klemick, H., Massey, D. M., Moore, C., Newbold, S., Simpson, R. D., ..... Wheeler, W. (2012). EPA valuation of surface water quality improvements. *Review of Environmental Economics and Policy, 6*(1), 130–146.

Gupta, A. K., Gupta, S. K., & Patil, R. S. (2003). A Comparison of Water quality indices for coastal water. *Journal of Environmental Science and Health, Part A, 38*(11), 2711–2725.

Horton, R. K. (1965). An index-number system for rating water quality. *Journal of the Water Pollution Control Federation, 37*(3), 300–305.

Hoyer, M. V., Brown, C. D., & Canfield, D. E., Jr. (2004). Relation between water chemistry and water quality as defined by lake users in Florida. *Lake and Reservoir Management, 20*, 240–248.

Jeon, Y., Herriges, J. A., Kling, C. L., & Downing, J. (2005). *The role of water quality perceptions in modeling lake recreation demand*. Iowa State University, Department of Economics Working Paper #05032.

Johnston, R. J., Besedin, E. Y., Iovanna, R., Miller, C. J., Wardwell, R. F., & Ranson, M. H. (2005). Systematic variation in willingness to pay for aquatic resource improvements and implications for benefit transfer: a meta-analysis. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie, 53*(2–3), 221–248.

Johnston, R. J., Besedin, E. Y., & Wardwell, R. F. (2003). Modeling relationships between use and nonuse values for surface water quality: a meta-analysis. *Water Resources Research, 39*(12), 1363.

Johnston, R. J., Schultz, E. T., Segerson, K., Besedin, E. Y., & Ramachandran, M. (2012). Enhancing the content validity of stated preference valuation: the structure and function of ecological indicators. *Land Economics, 88*, 102–120.

Khan, F., Husain, T., & Lumb, A. (2003). Water quality evaluation and trend analysis in selected watersheds of the atlantic region of Canada. *Environmental Monitoring and Assessment, 88*(1), 221–248.

Kung, H., Ying, L., & Liu, Y. (1992). A complementary tool to water quality index: fuzzy clustering analysis. *JAWRA Journal of the American Water Resources Association 28*(3), 525–533.

Landwehr, J. M., & Deininger, R. A. (1976). A comparison of several water quality indices. *Journal of the Water Pollution Control Federation, 48*(5), 954–958.

Liou, S. M., Lo, S. L., & Wang, S. H. (2004). A generalized water quality index for Taiwan. *Environmental Monitoring and Assessment, 96*(1), 35–52.

McClelland, N. I. (1974). *Water quality index application in the Kansas river basin*. EPA-907/9-74-001. Kansas City, MO: US EPA Region VII.

Mercer, P. (2003). Refined arithmetic, geometric, and harmonic mean inequalities. *Rocky Mountain Journal of Mathematics, 33*(4), 1459–1464.

Mitchell, R. C., & R. T. Carson. (1981). *An experiment in determining willingness to pay for national water quality improvements, report to the US Environmental Protection Agency*. Washington, D.C: Resources for the Future.

Mitchell, R. C., & Carson, R. T. (1989). *Using surveys to value public goods: the contingent valuation method*. Washington, D.C.: Resources for the Future.

Morgan, M. G., & Henrion, M. (1990). *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge, UK: Cambridge University Press.

Michael, H. J., Boyle, K. J., & Bouchard, R. (2000). Does the measurement of environmental quality affect implicit prices estimated from hedonic models? *Land Economics, 76*(2), 283–298.

Nagels, J. W., Davies-Colley, R. J., & Smith, D. G. (2001). A water quality index for contact recreation in New Zealand. *Water Science and Technology, 43*(5), 285–292.

Ott, W. R. (1978). *Water quality indices: a survey of indices used in the United States* (pp. 1–138). Washington DC: US Environmental Protection Agency. EPA-600/4-78-005.

Parparov, A., Hambright, K., Hakanson, L., & Ostapenia, A. (2006). Water quality quantification: basics and implementation. *Hydrobiologia, 560*(1), 227–237.

Pendleton, L., Martin, N., & Webster, D. G. (2001). Public perceptions of environmental quality: a survey study of beach use and perceptions in Los Angeles county. *Marine Pollution Bulletin, 42*, 1155–1160.

Pesce, S. F., & Wunderlin, D. A. (2000). Use of water quality indices to verify the impact of Córdoba City (Argentina) on Suquía River. *Water Research, 34*(11), 2915–2926.

Prakirake, C., Chaiprasert, P., & Tripetchkul, S. (2009). Development of specific water quality index for water quality supply in Thailand. *Songklanakarian Journal of Science and Technology, 31*(1), 91–104.

Sarkar, C., & Abbasi, S. A. (2006). Qualidex – a new software for generating water quality indice. *Environmental Monitoring and Assessment, 119*(1), 201–231

Simões, F. S., Moreira, A. B., Bisinoti, M. C., Gimenez, S. M. N., & Yabe, M. J. S. (2008). Water quality index as a simple indicator of aquaculture effects on aquatic bodies. *Ecological Indicators, 8*(5), 476–484.

Smith, D. G. (1990). A better water quality indexing system for rivers and streams. *Water Research 24*(10), 1237–1244.

Smith, V. K., & Desvousges, W. H. (1986). *Measuring water quality benefits*. Boston: Kluwer Nijhoff.

Swamee, P. K., & Tyagi, A. (2000). Describing water quality with aggregate index. *Journal of Environmental Engineering, 126*(5), 451–455.

Taner, M. Ü., Üstün, B., & Erdinçler, A. (2011). A simple tool for the assessment of water quality in polluted lagoon systems: a case study for Küçükçekmece Lagoon, Turkey. *Ecological Indicators, 11*(2), 749–756.

U.S. EPA. (2000). *A benefits assessment of water pollution control programs since 1972: Part 1, the benefits of point source controls for conventional pollutants in rivers and streams.*, Washington, DC: Final report to the U.S. EPA, Office of Water. EPA-68-C6-0021.

U.S. EPA. (2003a). *Environmental and economic benefit analysis of the final revisions to the national pollutant discharge elimination system regulation and the effluent guidelines for concentrated animal feeding operations*. Washington, DC: Office of Water. EPA-821-R-03-003.

U.S. EPA. (2003b). *Estimation of national economic benefits using the national water pollution control assessment model to evaluate regulatory options for Concentrated Animal Feeding Operations (CAFOs)*. Washington, DC: Office of Water. EPA-821-R-03-009.

U.S. EPA. (2004a). *Economic and environmental benefits analysis of the final effluent limitations guidelines and new source performance standards for the concentrated aquatic animal production IndU.S.try point source category.* Washington, DC: Office of Water. EPA-821-R-04-013.

U.S. EPA. (2004b). *Economic and environmental benefits analysis of the final meat and poultry products rule*. Washington, DC: Office of Water. EPA-821-R-04-010.

U.S. EPA. (2009). *Environmental impact and benefits assessment for final effluent guidelines and standards for the construction and development category*. Washington, DC: EPA Office of Water. EPA-821-R-09-012.

U.S. EPA. (2010). *Economic analysis of final water quality standards for nutrients for lakes and flowing waters in Florida*. Washington, DC: Office of Water.

Van Houtven, G., Powers, J., & Pattanayak, S. K. (2007). Valuing water quality improvements in the united states using meta-analysis: is the glass half-full or half-empty for national policy analysis? *Resource and Energy Economics, 29*(3), 206–228.

Van Houtven, G., Pattanayak, S. K., Patil, S., & Depro, B. (2011). Benefits transfer of a third kind: an examination of structural benefits transfer. In: J. Whitehead, T. Haab, & J.-C. Huang (Eds.), *Preference data for environmental valuation: combining revealed and stated approaches* (pp. 303–321). New York: Routledge.

Vaughan, W. J. (1981). The water quality ladder. In: R. C. Mitchell, & R. T. Carson (Appendix II), *An experiment in determining willingness to pay for national water quality improvement, draft report*. (Available at: http://yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0011-04.pdf/$file/EE-0011-04.pdf).

Walski, T. M., & Parker, F. L. (1974). Consumers water quality index. *Journal of the Environmental Engineering Division, 100*(3), 593–611.