Statistical considerations in the development of environmental indices: the example of the Canadian Water Quality Index

Dr. Sylvia R, Esterby[1] and Dr. Abdel H, El-Shaarawi[2]. [1]Mathematics, Statistics and Physics, University of British Columbia Okanagan, Kelowna, British Columbia, Canada, V1V 1V7 and [2]National Water Research Institute, and McMaster University, Burlington, Ontario, Canada, L7R 4A6.

Abstract: It is now widely accepted that human health and well-being are highly dependent on the quality of the environment, yet the health of the environment plays a much less prominent role in public policy than do economic considerations. The development of environmental indicators and aggregated indices is seen as a way to facilitate the integration of environmental information with social and economic information. Indicators constructed for public policy purposes are expected to provide information in a form readily understood by decision makers and the public. An indicator should also be quantitative and based upon an underlying model which relates the indicator to the more complex phenomenon being summarized. The Canadian Water Quality Index (WQI) provides an example of the development of an index where 1) it was recognized that simpler summaries were needed for the public and decision makers, 2) monitoring data were already being collected, and 3) the scientific understanding of water quality was used to guide the development of the index. During the development of an index and subsequent evaluation, statistical thinking and methodology can contribute in various ways. Since the WQI is an aggregated index, investigation of the value of the index, generated by various scenarios of underlying water quality variable values, needed to be assessed to ensure that low or high water quality according to the index agrees with what a more multivariate assessment would have given. A measure of variability of the index is also necessary. A number of issues, including these two, will be considered for the WQI using data from several Canadian provinces. The role and form of the WQI in the context of a suite of sustainability indicators will also be considered.

Keywords: water quality index ; sustainability; statistical consideration; public policy.

It is now generally agreed that the health of the aquatic environment should be monitored and that the results should be reported to decision-makers and the public, in addition to dissemination through technical publications. Some have contended that the way to provide this information to decision-makers and the public is to reduce complex information to a simple aggregate index which can be easily understood. The question arises as to whether such a reduction can be justified and under what conditions, since even a simple message must be defensible and accurate in the most important parts of the complex picture. In this paper we review some cases of the use of environmental quality indicators, consider criteria such indicators should meet, and discuss the Canadian Council of Ministers of the Environment (CCME) Water Quality Index (WQI).

Some examples of indicators and indices

Four examples which vary in some aspect are the following : 1) an indicator which measures quality relative to a specific use such as an indicator organism for safe drinking water, 2) a measurable variable used to trace the response of a water body to a regulation, where a decrease in ambient concentration was expected as the result in a decrease in loading (phosphorous), 3) a biological indicator as an integrator of effects of pollution on the health of the water body , and 4) a water quality index which is an aggregated index designed to represent some empirical model of reality.

The first three examples draw on North American experience. Microbiological quality is determined by testing drinking water for *Escherichia coli* (Health Canada 2010) which serves as the indicator organism for disease-causing bacteria, protozoa and viruses. Coliforms will be the first bacteria present in contaminated water and in larger concentrations than other types of pathogenic microbes, thus providing an indicator that is also readily detectable. A statistical consideration of drinking water regulations is given by Carbonez, El-Shaarawi and Teugels (1999). By 1964 extensive algal growth in the North American Great Lakes prompted the International Joint Commission, through participating US and Canadian agencies, to seek a means of remedying the situation (Stevens and Neilson, 1987). Phosphorus was deemed the limiting nutrient and regulations were enacted to control the loading of total phosphorus (TP) to Lake Ontario. To follow the response of the lake the 1972 and 1978 Great Lakes Water Quality Agreement s called for monitoring of ambient TP concentrations in Lake Ontario to track the response to reduced loading. An International Joint Commission Workshop was held in 1985 where scientists who specialized in the study or aquatic organisms debated the merits of using specific organisms as integrative indicators of pollution, where considerations such as broad occurrence of a species across the continent, sensitivity to pollutants were considered (Esterby, 1988).

The fourth example can be exemplified by the eutrophication index, one of six indicators comprising a composite pollution index for the Netherlands (Hammond *et al.*, 1995). Phosphates and nitrogen compounds are the primary substances contributing to eutrophication (although noted above, phosphorous compounds were thought to be limiting). Releases of phosphates and nitrogen compounds are expressed in eutrophication equivalents. Thus the aggregated index combines information on the eutrophication causing contributions as well as amounts into account. Thus the index quantifies a measure related directly to the problem identified, ie. eutrophication.

Perhaps the broadest environmental indicator is the EPI which in 2010 report ranked 163 countries on the basis of EPI score, itself built up from sets of indicators in categories (Emerson *et al*., 2010). Therein, the water quality index is a proximity-to-target composite indicator based on 5 water quality variables and with a station density adjustment. The current development of sustainability indicators for Canada involves 6 indicators (Government of Canada, 2007, and Sustainability Report, 2011), one of which is the CCME WQI, which will be considered in the remainder of the paper.

Criteria for aggregated indices

The report of the World Resources Institute (Hammond *et al*., 1995) considers what is required of an indicator that is to be used to communicate information to decision makers. The authors contend that we must have an underlying model of reality, the indicator must provide a metric against which some aspects of public policy can be measured, it must be analytically sound and have a fixed methodology of measurement, and we must be explicit about the model and metric.

It is clear that an aggregated index must be based on a model, since we must understand just what is being measured. The WRI report cites the example of the Dutch indicator for environmental eutrophication, mentioned above. Similarly, the Dutch indicator for climate change is derived from a weighted sum of the discharge of gases based on gases which increase the atmosphere’s warming potential. In both cases a phenomenon is identified and a metric based on measurable variables is defined.

Arguments for the role of single variable indicators for regulation versus indices for communication and broad comparisons could also be developed. When the action taken on the basis of an indicator is immediate and involves, for example, whether we will get sick if the organism is there or the contaminant has been dumped, we will tend to emphasize the technical validity of the indicator. When we must communicate environmental quality and make comparisons between diverse regions or countries, cases where there are a many variables which don't match from region to region or country to country, then we may have no choice but to use a broader brush, and aggregate and approximate more. The tool to use and the confidence we can attach will differ in these different scenarios. The case if diversity indices and progress towards the Convention on Biological Diversity also has provides insight on aggregated indices (Walpole *et al., 2009).*

Canadian water quality index

The Canadian Council of Ministers of the Environment (CCME) endorsed what is know as the Water Quality Index (WQI) in 2001 (Canadian Council of Ministers, 2001). It is the result of a joint effort of the federal and provincial Canadian governments. The index is basically the index developed in the 1990s by the province of British Columbia and modified to match the expectations of the Alberta index such that high values of the index correspond to good quality.

Two types of data are required to calculate the WQI. First, measurements on several water quality variables made over the period and region of interest. Second, numerical values representing water quality objectives associated with each water quality variable. Three measures of violations are computed, called factors, based on comparing the data with their objectives and the factors are combined to provide the value of the index. The factors are:

F1, the Scope equal to the percentage of variables that failed to meet their objectives (limits) at least once during the time period under consideration.

F2, the Frequency equal to the percentage of individual tests that failed to not meet objectives.

F3, the Amplitude which measures the amount by which each failed observation differs from its objective.

The ideal data can be represented as a matrix **X** with p columns and n rows where the entry in i^{th} row and j^{th} column x_{ij} represents the measurement on j^{th} water quality variable (j =1, 2, …, p) for i^{th} case (i =1, 2, …, n). For example if the index was being computed for a sampling station, where monthly samples were taken for one year, i=1,2,...,12. There is one or more water quality limit(s) u_j associated with the j^{th} variable (column). It is more common to have the objective stated as a single limit representing an upper or lower concentration limit for the water quality variable. An example of an upper limit objective is bacterial level (E coli in drinking or recreational waters) and an example of a lower limit is dissolved oxygen concentration in the water. In other cases an allowable range is used as in the case of pH or when dealing with water level in the cases of draughts and floods. The index is defined as

$$WQI = 100 - \frac{\sqrt{F_1^2 + F_2^2 + F_3^2}}{1.732}$$

Where:

$$F_1 = \frac{Number\ of\ \ failed\ \mathrm{var}iables}{Total\ number\ of\ \ \mathrm{var}iables} \times 100$$

$$F_2 = \frac{Number\ of\ \ failed\ tests}{Total\ number\ of\ tests} \times 100$$

$$F_3 = \frac{nse}{0.01nse + 0.01}$$

$$nse = \frac{\sum_{i=1}^{n} excursion_i}{\#of\ Tests}$$ and excursions are defined as

$$excursion_i = \left[\frac{Failed\ \ test\ value_i}{objective_j}\right] - 1$$ when the failing observed value exceeds the objective

and $excursion_i = \left[\frac{objective_j}{Failed\ \ test\ value_i}\right] - 1$ when the failing observed value is less than the objective.

Some Shortcomings of the WQI Formulation

There are a number of difficulties with the index formulation which can be seen by examination of the definitions above.

1. It concentrates on the violations of the criteria and thus ignores the data below the threshold. The ignored data contain information about trends and not including such information in the index means that the index will have low trend detection power. It is possible that the values are in compliance with the objectives but they are progressing toward non-compliance and such an index should reflect this trend tendency.

2. It is not easy to assess the statistical uncertainty in the index since the three factors F_1, F_2 and F_3 are not independent and indeed they are positively dependent.

3. The distribution of non-compliant values among the variables will have a major impact on the value of F_1. This can be seen when the number of non-compliant values are equal to the number of variables p. In this case the possible values of F_1 will range between 100/p and 100 with the minimum value of F_1 occurring when all the p values occur in a single variable while its maximum occurs when a single non-compliance occurs in each of p variables.

4. The index does not allow for weights to reflect the importance of the variables. Clearly different pollutants should not be symmetrically treated.

5. The division by the observed value when computing nse or variables which fail compliance for values below a lower limit is inappropriate. The proportional exceedance should always be relative to the objective. To make it relative to the observed value means that the comparison changes with each different observed value and this makes F_3 unstable.

6. The Index as a statistic does not have a measure of uncertainty.

Points 3 and 4 above imply the exchangeability of variables. A particular value of the index that arises from one water quality variable being responsible for all the failures and another where more than one variable failed at least one time, under suitable values for the excursions, would be considered to have equal water quality. Clearly, this loss of information would be unsatisfactory on many occasions, and for example make use of WQI to assess changes over time meaningless.

Numerous reports are now available on application of the WQI (eg. CCME National Water Quality Index Workshop, 2004; Khan, Paterson and Khan, 2004; Mercier *et al,* 2005; Kilgour and Associates, 2009; Gartner Lee Limited, 2006) and often a caution is given that comparisons between regions cannot be made.

A way forward

There seems to be one major oversight in the use of the WQI. Returning to the criteria set out by the WRI report and particularly the requirement that a model be conceived at the outset, the question arises as to just what it is the WQI is measuring. It seems important to decide why it is being used more carefully. What is the environmental problem being tracked. For example, is eutrophication the problem? Or is it a problem of persistent toxic contaminants in the water, and so on. Once this is done, the variables of importance can be determined and taken as a fixed set. As long as the guidelines used are stated and used in all regions to be compared, and, of course held fixed if to be followed over time, then some of the problems with definition are eliminated. At least one of the reasons for the qualifiers on its use will be removed. There is merit in the original ideas of considering the number of failures and how far the failing observations are from the objectives. Further examination of how the factors F1, F2 and F3 can be used to characterize water quality are worth considering.

References

Canadian Council of Ministers, 2001. Canadian water quality guidelines for the protection of aquatic life, CCME Water Quality Index 1.0, Technical Report. 13pp.

Carbonez, A., El-Shaarawi, A. H. and Teugels, J. L. 1999. Maximum microbiological contaminant levels. Environmetrics, 10, 79-86.

CCME National Water Quality Index Workshop. 2004. A path forward for consistent implementation and reporting, Nov 24-25, 2003, Halifax, Nova Scotia, Workshop Proceedings.

Emerson, J., Esty, D., Levy, M.A., Kim, C.H., Mara, V., de Sherbinin, A. and Srebotnjak, T. 2010. Environmental Performance Index. New Haven: Yale Center for Environmental Law and Policy.

Esterby, S.R. 1988. Statistical methods for studying biotic responses to persistent toxic substances. In: Toxic Contaminants and Ecosystem Health; A Great Lakes Focus, M.S. Evans (Ed.), Advances in Environmental Science and Technology, Vol. 21, Wiley, pp. 447-475.

Gartner Lee Limited. 2006. A sensitivity analysis of the Canadian Water Quality Index. PN 1355. Report to Canadian Council of Ministers of the Environment. 36pp.

Government of Canada. 2007. Canadian Environmental Sustainability Indicators. Environment Canada, Statistics Canada and Health Canada. Cat no. 16-251-XIE.

Government of Canada. 2011. The Sustainability Report, Canada's Sustainability Indicators Initiative. http://www.sustreport.org/indicators/nrtee_esdi.html

Hammond, A., Adriaanse, A., Rodenburg, R. Bryant, D. And Woodward, R. 1995. Environmental Indicators: A systematic approach to measuring and reporting on environmental policy performance in the context of sustainable development. World Resources Institute. ISBN 1-56973-026-1.

Khan, A.A>, Paterson, R. And Khan, K. 2004. Modification and Application of the CCME WQI for the communication of drinking water quality in Newfoundland and Labrador. Water Qual. Res. J. Canada, 39, 285-293.

Kilgour & Associates Ltd. 2009. Reducing the sensitivity of the water quality index to episodic events. PN1435. Report to Canadian Council of Ministers of the Environment. 23pp.

Mercier, V. and 6 co-authors, 2005. Application and testing of the water quality index in Atlantic Canada, Report Summary. CCME web page. http://www.ccme.ca/assets/pdf/awi_en_posting.pdf

The Sustainability Report. 2011. Guidelines for Canadian Drinking Water Quality, http://www.hc-sc.gc.ca/ewh-semt/pubs/water-eau/2010-sum_guide-res_recom/index-eng.php#a2

Stevens, R. J. J., and Neilson, M. A. 1987. Response of Lake Ontario to reductions in phosphorus load, 1967-82. Can. J. Fish. Aquati.Sci., 44, 2059-2068.

Walpole, M., and 24 co-authors. 2009. Tracking progress toward the 2010 biodiversity target and beyond. Science, 325, 1503-1504.