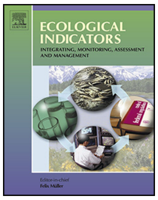
Ecological Indicators 45 (2014) 494–507

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind



The method matters: A guide for indicator aggregation in ecological assessments

Simone D. Langhans*, Peter Reichert, Nele Schuwirth

Eawag: Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland

ARTICLE INFO

Article history:
Received 12 January 2014
Received in revised form 14 May 2014
Accepted 16 May 2014

Keywords:
Aggregation
Multi-criteria decision analysis
Multi-attribute value theory
One-out all-out
River assessment
Water Framework Directive

ABSTRACT

Ecological assessment requires the integration of many physical, chemical, and/or biological quality elements. The choice of the aggregation method of such partial assessments into an overall assessment can considerably affect the assessment outcome – an issue that has been controversially discussed within the scientific community for the last decade. Current practice often considers only two different aggregation methods, the weighted arithmetic mean (additive aggregation) and the one-out, all-out method (minimum aggregation). However, both have important drawbacks. Additive aggregation compensates a bad status of one quality element by a number of elements featuring good status. Minimum aggregation can lead to overly pessimistic assessment results, since only the quality element in the worst status is considered. Here, we introduce a toolbox containing current and new aggregation methods, demonstrate and discuss their properties with simple, didactical examples, and suggest in which situations best to use them. Then, we illustrate the consequences of selected aggregation schemes for ecological river assessment with the case study of the Swiss Modular Concept of stream assessment (SMC), which we apply to ten river reaches in the Mönchaltdorfer Aa catchment in Switzerland. To be able to do so, we used multi-criteria decision analysis, i.e., multi-attribute value theory, to arrange the SMC quality elements into an objectives hierarchy, and to translate their individual assessments into value functions. Our case study revealed that choosing the most appropriate aggregation method particularly matters, if objectives with significantly different qualities are aggregated. We argue that redundant objectives (i.e., quality elements), often found at the lower levels of the objectives hierarchy, should best be aggregated additively allowing for compensation to increase the statistical significance of the results. Further, we suggest that complementary sub-objectives that often occur at higher levels may be optimally aggregated with a mixture of additive and minimum aggregation. Such a mixed method will allow some compensation, but nevertheless penalize for very bad states. Since here we compare commonly used aggregation methods with some which we believe have never been discussed in an assessment context before, our study concurrently informs ecological assessment in theory and in practice.

1. Introduction

Over the last few decades, freshwaters have suffered from a multitude of pressures resulting in poor ecosystem condition and a drastic decrease in biodiversity (Dudgeon et al., 2006; Vörösmarty et al., 2010). In an attempt to address these problems, a key management issue is to assess the ecological status of freshwater ecosystems, to identify the main perturbations responsible for the observed condition, and to put regulations and recommendations into place for ecosystem recovery (e.g., the EU Water Framework Directive (WFD); European Commission, 2000). Thereby, to provide a balanced view of the ecological status of freshwaters, evidence from multiple biological, chemical, physical and hydrological quality elements is usually combined (Moss et al., 2003).

Examples of such comprehensive ecological river assessment schemes are widespread including for instance the WFD's Common Implementation Strategy (European Commission, 2003), the Ecosystem Health Monitoring Program in South East Queensland (Bunn et al., 2010), the National Rivers and Streams Assessment (USEPA, 2013), and the Swiss Modular Concept for Stream Assessment (SMC; Bundi et al., 2000). Although these programmes differ in the selection of indicators and the spatial and temporal monitoring scheme, all of them integrate different quality elements

---

* Corresponding author. Current address: Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Müggelseedamm 310, 12587 Berlin, Germany.
Tel.: +49 0 30 64181 618; fax: +49 0 30 64181 750.
E-mail addresses: langhans@igb-berlin.de (S.D. Langhans), reichert@eawag.ch (P. Reichert), nele.schuwirth@eawag.ch (N. Schuwirth).

http://dx.doi.org/10.1016/j.ecolind.2014.05.014
1470-160X/

to higher-level indices, and some even to an overall score (Bunn et al., 2010; European Commission, 2000). This integration is usually done by using the one-out, all-out rule (hereafter referred to as the minimum aggregation; (European Commission, 2005; USEPA, 2013)), by averaging (hereafter referred to as the arithmetic mean or additive aggregation; (Barbour et al., 1999; Plafkin et al., 1989)), or by using one or the other at different hierarchical aggregation levels (LAWA, 2002; Smith and Storey, 2001).

The logic behind the application of a minimum aggregation in ecological assessment is that a river should not reach a good ecological status if any of the quality elements measured fail. This precautionary principle might be an appropriate approach for serious impacts, such as for instance a toxic level of a hazardous substance. However, for less acute pressures (Moss, 2007) the minimum aggregation increases the likelihood that we report a lower quality than the actual ecological status (Hering et al., 2010; Sandin, 2005), which is referred to as the pessimism bias (Cunningham, 2012). This source of pessimism is amplified by the number of quality elements included (Heiskanen et al., 2004). The additive aggregation, on the other hand, implies that a low value of one quality element can be compensated by large values of other quality elements. Therefore, it poses the risk of overlooking an impact, which in fact would ask for a measure.

Although a range of alternative aggregation methods exists, ecological river assessments have rarely adopted other methods than minimum aggregation or averaging. The reason for this may be the lack of studies that quantify the consequences, such alternative aggregation methods may have on the quality evaluation of comprehensive river assessment schemes.

Quantifying the ecological state of a river calls for a framework that allows assessing different elements of the river ecosystem, and aggregating these assessments to an overall score. Multi-criteria decision analysis (MCDA), specifically multi-attribute value theory (MAVT) (Eisenführ et al., 2010; Keeney, 1982; Keeney and Raiffa, 1976), offers such a framework (Corsair et al., 2009; Klauer et al., 2006; Reichert et al., 2007). In this framework, a value function represents the degree of fulfilment of the overall- or sub-objectives on a scale from zero to unity as a function of objectively measurable system properties, the attributes. In the river assessment terminology, the term attributes refers to indicators or assets (Langhans et al., 2013). To facilitate the construction of such a value function, the overall objective (aka the goal of the assessment) is broken down hierarchically in complementary sub-objectives (often referred to as assessment endpoints or quality elements in the river assessment terminology) that make the higher-level objective more concrete. The value function for the overall objective is then constructed by formulating individual value functions for each lowest-level sub-objective, as a function of a small number of attributes, and aggregating the values at higher levels. This requires the specification of value functions for all lowest-level objectives and aggregation rules at all higher levels.

If the overall objective is to reach a good ecological state of the river, the corresponding value function reflects an ecological assessment score. In other words, it quantifies the degree to which the good ecological state of the river is reached (Langhans et al., 2013). Similarly, sub-objectives describe the state of sub-systems, such as for example the invertebrate community or water quality (Fig. 1). This makes it possible to use such a value function informally for deficit analysis, as it is mostly done with traditional river assessment procedures, or to include it into a formal decision support process in environmental management.

The main objective of this study was to provide a toolbox containing a mix of currently applied and new aggregation methods along with some guidance on which one best to select. Hence in the following, we introduce a sequence of four generic aggregation methods that span the spectrum from allowing for full compensation of poor assessments of sub-systems to no compensation at all. In addition to these four generic types, we establish a range of alternative methods to allow for a finer resolution of the adequate degree of compensation. We then derive important properties of the aggregation methods, and investigate how they can affect classification outcomes. To do so, we compared hypothetical examples and a monitoring dataset from 10 river reaches in Switzerland assessed according to the Swiss Modular Concept of stream assessment (SMC). Thereby, we used MAVT to arrange the different SMC-quality elements in an objectives hierarchy and to translate their individual assessments into value functions. User guidance for the different aggregation methods was developed considering the properties as well as the on-ground assessment outcomes.

2. Material and methods

Aggregation methods integrate the values (which are the degrees of fulfilment of sub-objectives in decision science), v_i, to an overall value, v, representing the degree of fulfilment of the higher-level objective. An aggregation method is defined as a function f: $v=f(v_1,v_2,\ldots,v_n)$ that specifies how the higher-level value is calculated from the n values at the lower level. If all the sub-objectives are fulfilled to the same degree, it seems reasonable to assume that the higher-level objective is fulfilled to the same degree. This leads to the following condition for the aggregation function f:

$$f(v_1 = v, v_2 = v, \ldots, v_n = v) = v \tag{1}$$

In this paper, we will only consider aggregation methods that fulfil this condition.

2.1. Basic aggregation methods

To start off, we considered four generic aggregation methods that are either widely applied in river assessment (the weighted arithmetic mean (Eq. (3) and the minimum aggregation (Eq. (6)), or are rarely considered, but belong to the three most prominent means (the weighted geometric mean (Eq. (4) and the weighted harmonic mean (Eq. (5)). Note that for the aggregation methods (3)–(5), we assumed that the weights are normalized to sum up to one:

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

2.1.1. The weighted arithmetic mean (hereafter called additive aggregation)

For additive aggregation, the aggregated value is calculated as the sum of the n values, v_i, of the sub-objectives each of them multiplied with its weight, w_i:

$$f_{\text{add}}(v_1, \ldots, v_n) = \sum_{i=1}^{n} w_i v_i = w_1 v_1 + w_2 v_2 + \cdots + w_n v_n \tag{3}$$

If the weights are equal for all elements ($w_i = 1/n$), the result is identical to the (unweighted) arithmetic mean which is often referred to as unweighted averaging (Guitouni and Martel, 1998). In decision science, the weighted arithmetic mean is called additive aggregation, which is by far the most widely used aggregation function for multi-criteria decision support (Keeney and Raiffa, 1976; Eisenführ et al., 2010).