# Towards an integrated surface water quality assessment: Aggregation over multiple pollutants and time

Nele Schuwirth

*Eawag: Swiss Federal Institute of Aquatic Science and Technology, Ueberlandstrasse 133, 8600, Dübendorf, Switzerland*

## A R T I C L E   I N F O

## A B S T R A C T

Surface water quality management requires foresighted decision making regarding long-term investments. It should consider multiple objectives (e.g. related to different pollutants and costs), integrate multiple sources of pollution (point and diffuse sources), and external conditions that change over time (climate, population and land-use changes). Multi-attribute value theory can support such decisions, especially the development of an assessment method. Integrated surface water quality assessment methods including micropollutants are currently lacking or in development in many countries. Important steps for the development of such an immission oriented and integrated surface water quality assessment method are discussed in this paper and exemplified for organic micropollutants. The proposed assessment method goes beyond simple pass-fail criteria for single substances. It provides a continuous assessment on a scale from zero to one based on five color-coded water quality classes and suggestions for the visualization of assessment results. It takes into account the toxicity of the micropollutants and their mixture to aquatic organisms by comparing measured concentrations to environmental quality standards (EQS). The focus of this paper is on aggregation over multiple substances and time. Advantages and disadvantages of different aggregation methods are discussed as well as their implications for practice. The consequences of different aggregation methods are illustrated with didactical examples and by an application of the proposed water quality assessment method to pesticide monitoring data from Switzerland. Recommendations are provided that account for the purpose of the assessment. Furthermore, the paper illustrates how the proposed method can facilitate dealing with uncertainty and a transparent communication of monitoring results to support water quality management decisions.

## 1. Introduction

Micropollutants (e.g. pesticides, pharmaceuticals) are increasingly recognized as a threat to freshwater ecosystems (Schwarzenbach et al., 2006) and many countries set up monitoring programs for micropollutants in surface waters. Still, so far integrated assessment methods are lacking that translate individual measurements into an overall assessment of the water quality state, going beyond pass/fail criteria of single substances and single monitoring samples. This makes it difficult to assess the overall water quality and to follow spatial and temporal developments at various scales. An integrated assessment of the water quality of surface waters has to cover different substances (from nutrients to micropollutants) and different sources of pollution (diffuse and point sources) (Schwarzenbach et al., 2006). It has to account for

the toxicity of different substances and their degradation products, and for mixture toxicity (Drakvik et al., 2020). It should consider exposure patterns, since concentrations of many substances vary largely over time, especially those from diffuse sources such as agriculture or sewer overflows. Furthermore, the spatial variation of exposure within catchments depends on the position of point sources in the river network, land use patterns in the catchment, and the dilution in the river with increasing discharge.

In this paper, a framework based on multi-attribute value theory (Dyer and Sarin, 1979; Eisenfuehr et al., 2010) is introduced that is able to cover the different aspects and can deal with different sources of uncertainty. It follows an immission oriented approach to assess the surface water quality in the receiving waters at the site and catchment or river basin scale. It relies on environmental quality standards that are based on ecotoxicological hazard assessment and considers mixture toxicity. Such a procedure can hardly be comprehensive due to missing environmental quality standards for many substances and the challenge to

*E-mail address:* nele.schuwirth@eawag.ch
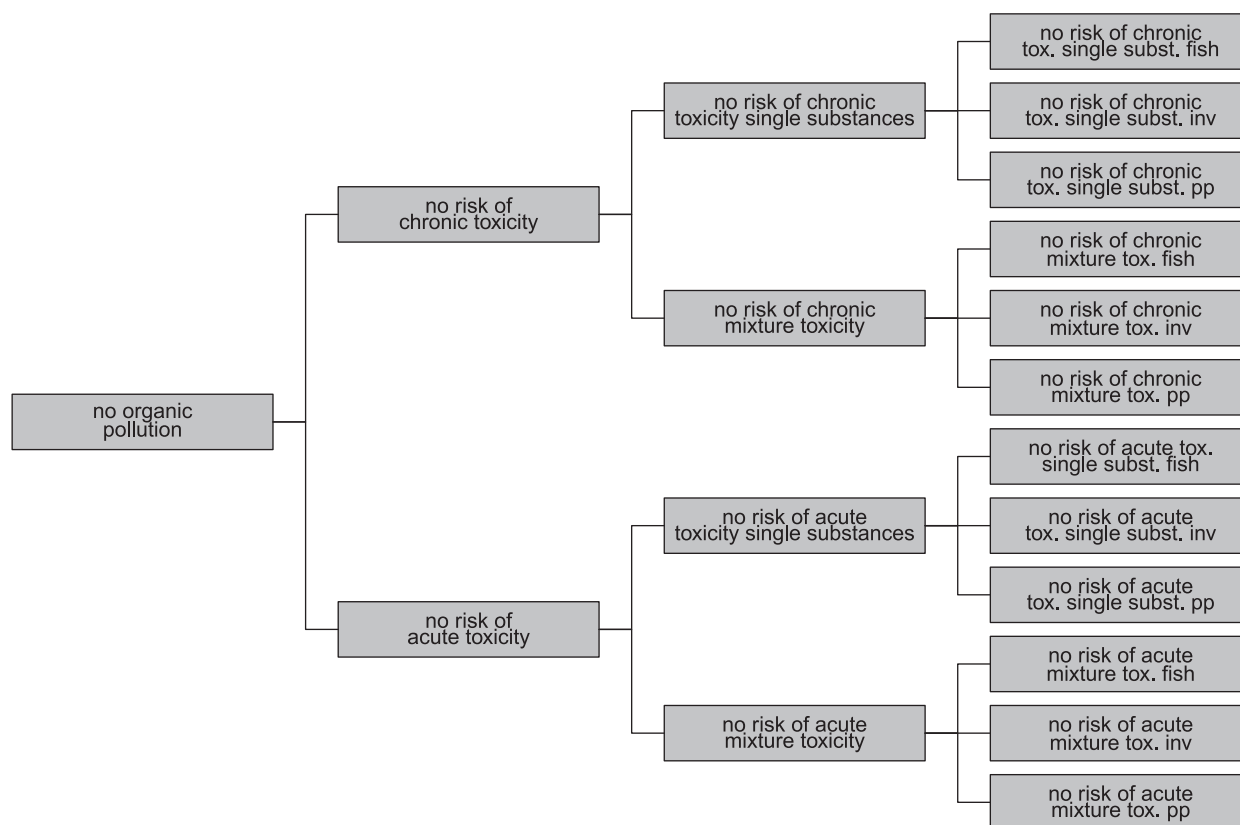
**Fig. 1.** Example for an objectives hierarchy for water quality assessment regarding organic pollution where the main objective (no organic pollution) is on the left, and the lower level objectives are towards the right.

quantify all substance concentrations. Therefore, it would be advisable to complement the proposed method with effect-based tools (biotests) and non-target screening of contaminants (Brack et al., 2017). However, this goes beyond the scope of this paper. Potentially controversial steps for the development of such an assessment and critical decisions that have to be taken based on expert knowledge are introduced in the following.

The development of an integrated water quality assessment method based on multi-attribute value theory can be divided into the following steps (Eisenfuehr et al., 2010):

1) Objectives hierarchy: The development of an objectives hierarchy facilitates a transparent and explicit formulation of all water quality aspects that should be assessed. It provides a structure that helps linking the overall objective (good water quality) to more concrete sub-objectives (e.g. low concentrations of specific substances) (see Fig. 1 for an example). The hierarchical structure supports the aggregation of multiple objectives into an overall objective (Eisenfuehr et al., 2010).

2) Measurable attributes: The lowest level sub-objectives are linked to measurable attributes (e.g. substance concentrations). In the scope of water quality assessment the development of an objectives hierarchy has to consider the selection of substances that have to be monitored and assessed as well as their toxicity to aquatic organisms and mixture toxicity (Junghans et al., 2013; Spycher et al., 2018).

3) Continuous value function: Common for the assessment of the ecological status of rivers in Europe are the definition of water quality classes that allow a color-coding of the assessment results and facilitate the communication of results: high (blue), good (green), moderate (yellow), poor (orange), bad (red), according to the Water Framework Directive (European Union, 2000). The same classification is applied to

the proposed water quality assessment method. However, to avoid discretization errors and to allow us to evaluate differences within a quality class, a continuous assessment scale from 0 (bad) to 1 (high) is applied that describes the fulfilment of the objectives and can be discretized into the water quality classes at each level of the objectives hierarchy (e.g. Langhans et al., 2013). Following multi-attribute value theory (Dyer and Sarin, 1979), we can construct such a value function for each objective at the lowest level of the objectives hierarchy that directly depends on measurable attributes and is not restricted to any functional form.

4) Hierarchical aggregation: To assess the fulfilment of objectives at higher levels of the objectives hierarchy, the values of the corresponding lower level nodes have to be aggregated with an aggregation rule that reflects how the fulfilment of the higher-level objective depends on the fulfilment of its sub-objectives (Langhans et al., 2014; Haag et al., 2019; Reichert et al., 2019). In principle, all functions from the family of so called quasi-arithmetic means can be considered (Grabisch et al., 2009).

5) Temporal aggregation: Exposure patterns of different substances vary largely over time and are affected by weather conditions, which have to be accounted for by an appropriate monitoring design (Wittmer et al., 2014). To assess the effects of management actions and to follow temporal trends, different temporal time scales may be of interest. For example, to assess the seasonal development within one year, one would be interested in the results at the resolution of the samples. To compare different sites or different years, it might be reasonable to aggregate over the samples within each year and site. In principle, the same aggregation functions can be used as for the hierarchical aggregation and the choice can be made based on a similar reasoning, as described below.

6) Spatial aggregation: Exposure patterns vary not only in time but also in space. The dendritic network structure of rivers, the hydrological conditions, and the location of point and non-point sources lead to variations of exposure patterns in space. Water quality monitoring was traditionally often focused on the outlet of a catchment, which allows a good overview about fluxes, but does not provide information about concentrations within the network. This can lead to the fact that an exceedance of toxic thresholds within the network is overlooked, especially in headwater streams with a low dilution of pollutants. Whether or not a spatial aggregation is useful depends on the purpose of the assessment and the spatial resolution of the monitoring program. It could be useful to quantify the water quality status of whole river basins or catchments, for example in the scope of prioritizing management measures within the catchment to maximize the overall ecological state of the catchment. Since water quality monitoring is typically restricted to a few points within a river network, this would usually require an inter- and extrapolation of monitoring results to the whole network, which should consider the location of point- and non-point sources of pollution. One of the fundamental objectives of immission-based surface water quality assessment is the quantification of the ecological status to protect the organisms inhabiting the ecosystems, such as fish, invertebrates and primary producers. For those organisms, the spatial arrangement of habitats matter and have to be considered to assess the ecological status of catchments as a whole (as opposed to the assessment of specific reaches or sites). Since water quality is an important aspect of habitat quality, the water quality assessment should ideally be integrated into an ecological assessment at the catchment scale (Kuemmerlen et al., 2019).

7) Propagation of uncertainty: Different sources of uncertainty influence the water quality assessment, from sampling, chemical analytics, environmental quality standards (Baudrot and Charles, 2019), to parameters of the value function including aggregation. Multi-attribute value theory offers a transparent framework that allows us to propagate and assess these uncertainties to evaluate the robustness of the final assessments (Schuwirth et al., 2012; Scholten et al., 2015; Schuwirth et al., 2018).

8) Communication of results: Communication of risks is an important step towards more comprehensive risk assessment (Topping et al., 2020). Monitoring and assessment usually serves multiple purposes and has to address different audiences. It should inform experts by allowing detailed analyses with high temporal and spatial resolution to identify deficits and potential management actions, and it should provide a synthesized overview to inform policy makers, stakeholders, or the public and to support the planning of management actions in other sectors, such as morphological river restoration. Therefore, it is important to provide visualizations and numerical results at different levels of the objectives hierarchy and at different temporal and spatial scales (Reichert et al., 2013; Schuwirth et al., 2018).

The focus of this paper is on the aggregation over multiple pollutants and time. A comprehensive coverage of spatial aggregation goes beyond the scope of this paper, because it requires the inter-/extrapolation of the attributes to the whole river network and consideration of additional ecological criteria. The paper provides recommendations on how to deal with these challenges by discussing each step in detail (Section 2), and an illustration of the proposed methods with existing pesticide monitoring data from Switzerland (Section 3). The aim of this paper is to support the development of integrative assessment methods for the water quality of surface waters.

## 2. Methods

I use multi-attribute value theory to formulate an integrated water quality assessment method following the eight steps introduced above (Dyer and Sarin, 1979) based on insights from discussions with a working group in Switzerland that aims for the development of a water quality assessment method with different stakeholders from cantonal and federal authorities.
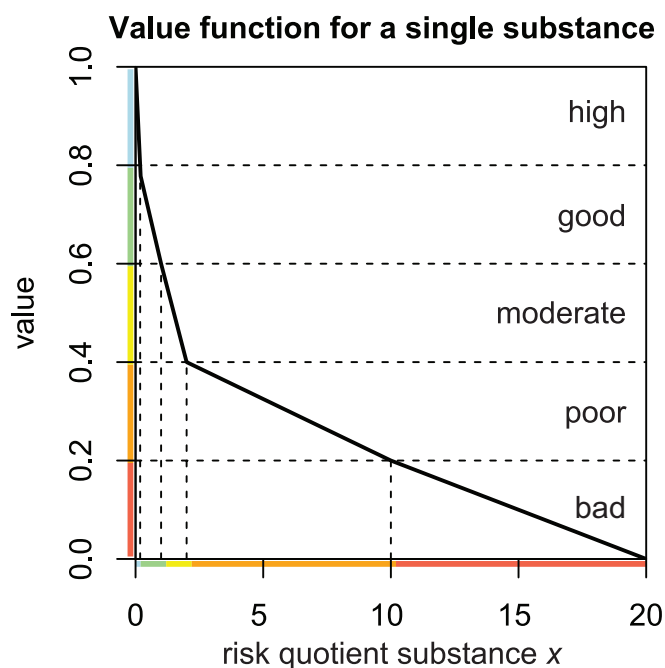
**Objectives** for water quality assessment include "no ecotoxicological risk" for different groups of aquatic organisms from single substances and from the mixture of substances. Acute and chronic effects should be considered. The **hierarchical structure** regarding these objectives should be chosen in order to facilitate the synthesis and communication of the results and support management. The objectives hierarchy should be comprehensive, and sub-objectives that are not in the same branch should be complementary. On the other hand, the hierarchy should be as concise as possible to facilitate the communication of results (Marttunen et al., 2019). Different structures are possible and should be discussed with stakeholders that are going to use the assessment method. One possible hierarchical structure breaks down the overall objective of good water quality into different modules that group the pollutants according to the sampling and measurement strategy (e.g. nutrients, organic micropollutants, heavy metals and metalloids). This ensures that the results of a measurement campaign can be presented together. Other structures are possible, for example to further sub-divide substance groups or to distinguish different sources of pollution. However, since many substances can have various sources, the latter is more difficult to implement. For organic micropollutants, we can distinguish chronic and acute toxicity, single substances and mixture toxicity (Fig. 1).

In addition to the ecotoxicological risk assessment, it might be advisable to include objectives that cover the environmental pollution with substances that are typically not of concern because of their toxicity, but still do not belong in a natural environment (Lienert et al., 2011), such as x-ray contrast agents, artificial sweeteners, or repellents. This could be implemented by an additional objective for the general prohibition of pollution by non-natural substances.

The next step is to identify measurable **attributes** that allow us to measure the fulfilment of the lowest level sub-objectives. The selection of substances to be monitored and sampling strategies should reflect the objectives and are crucial aspects of the monitoring design (Behmel et al., 2016; Norman et al., 2020). These aspects are not discussed here in all detail but they should be standardized to ensure the comparability of results within one monitoring program. It would make only limited sense to aggregate data from different sampling strategies.

Wittmer et al. (2014) propose a concept for the sampling and assessment of organic micropollutants in Switzerland, suggesting biweekly composite samples to be compared with chronic environmental quality standards. This design takes into account that micropollutants from non-point sources, such as agricultural pesticides, have seasonal exposure patterns that vary with application times and depend on weather conditions. A yearly averaging would lead to a strong dilution of the concentration of these substances, which would not be representative for the exposure situation of stream organisms. Ecotoxicological tests to assess the chronic toxicity usually have a duration between 72h and several weeks. A biweekly sampling strategy therefore roughly corresponds to the duration of the chronic toxicity tests (Wittmer et al., 2014). This complies with the new Swiss Waters Protection Ordinance[1]. For

---

[1] https://www.bafu.admin.ch/bafu/de/home/themen/wasser/mitteilungen.msg-id-78122.html (available in German, French and Italian)

## Value function for a single substance



**Fig. 2.** A continuous value function to assess the chemical status regarding a single substance based on its risk quotient (measured environmental concentration divided by the environmental quality standard). The same function can be applied to assess mixture toxicity based on the sum of risk quotients, see text. For risk quotients above 20 the value stays at zero; this is not shown here to improve readability.

the assessment of acute effects, 24-96h composite samples would be ideal. We can then use the concept of risk quotients, where the measured concentrations in the water samples are divided by the environmental quality standard.

Mixture toxicity can be assessed according to Junghans et al. (2013), where the risk quotient of substances that affect the same organism group are summed up (Spycher et al., 2018). As additional information, an assessment of the toxicity of single substances can be provided to support the identification of the sources of pollution and potential management measures. For this, the highest risk quotient of the single substances for each organism group can be used.

The following class boundaries were suggested to link the risk quotient of single substances or mixtures to five color-coded water quality classes (Wittmer et al., 2014): high (blue): RQ<0.1; good (green): RQ<1; moderate (yellow): RQ<2; poor (orange) RQ<10; bad (red) RQ≥10.

A straightforward way to derive a **continuous value function** from discrete water quality classes is to assume equally spaced intervals at the value scale and apply linear piecewise interpolation between class boundaries (e.g. Fig. 2), but any other functional form is possible that best reflects how the fulfilment of the objective depends on the attribute level.

Here, the class boundary between the moderate and the good state is based on the legal requirements (in this case the EQS valuei.e. risk quotient=1). Such a continuous assessment and its translation into five color-coded quality classes provides a graded system that can resolve differences between sites or temporal changes better than just a binary pass/fail criterion currently applied for the assessment of chemical contamination in the European Water Framework Directive (Brack et al., 2017).

Most commonly used aggregation techniques for **hierarchical** and **temporal aggregation** are *minimum aggregation* (i.e. worst-case or one-out all-out) and *additive aggregation* (weighted arithmetic mean) (see Table 1 for examples).

While *minimum aggregation* just takes the worst value of the sub-objectives as the aggregated value, the *additive aggregation* takes a (weighted) average across all sub-objectives, which means that it allows for full compensation between good and bad values. Many other less widely used aggregation techniques exist that provide a compromise between those extreme aggregation methods (Langhans et al., 2014; Haag et al., 2019; Reichert et al., 2019) (see Table 2 in the main text for examples and their mathematical equations).

To choose an appropriate aggregation function, it is important to be aware about the properties that the aggregation functions should have in the scope of the assessment. Properties for aggregation techniques that seem rather undisputed in the scope of water quality assessment are:

(1) If the values of all sub-objectives are the same, the aggregated value is also the same (idempotency).
(2) If the value of one of the sub-objectives improves, the aggregated value does not get worse (and vice versa).

These two properties may seem rather obvious. They are also mentioned as a precondition in mathematical aggregation theory and their application to sustainability assessment (Pollesch and Dale, 2015). Still, not all aggregation methods discussed in the literature fulfil them. For example, the multiplicative aggregation method introduced in multi-criteria decision making (Keeney and Raiffa, 1976) does not fulfil the idempotency requirement (1). And even in the field of water quality assessment, the aggregation methods proposed by Swamee and Tyagi (2000, 2007) do not fulfil the idempotency requirement. Furthermore, these two methods depend on the number of sub-objectives to be aggregated, which is another undesired property.

Further desirable properties that were mentioned in discussions with stakeholders are:

(3) The aggregation technique is easy to understand and communicate.
(4) The aggregation technique is sensitive to changes and facilitates the detection of improvements and deteriorations (in space or time, e.g. across years), i.e., if one of the sub-
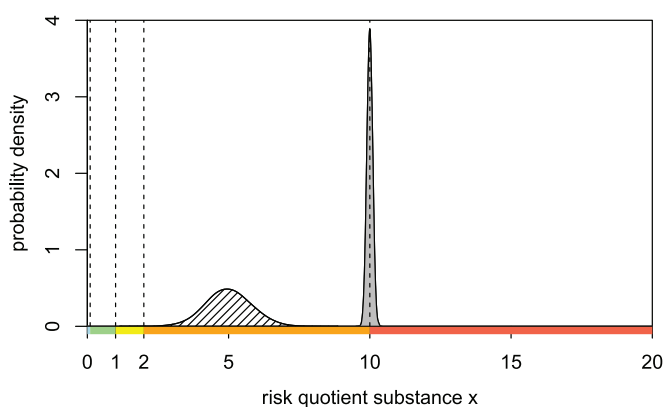


**Fig. 3.** Illustration of the connection between the uncertainty in the attribute value (here risk quotient of a substance x) and the associated uncertainty of the classification (color-coding). In this example, the grey colored distribution has an expected value of 10, which is at the boundary between the bad (orange) and poor (red) quality class. Even though the uncertainty about this value is rather low (normal distribution with a relative standard deviation of 1%), the uncertainty about the class is rather high (50% probability for the poor and 50% probability for the bad class). The shaded distribution has an expected value of 5, and despite its comparably large uncertainty (normal distribution with a relative standard deviation of 16%) the uncertainty regarding the classification is very low (99.99% probability for the poor class). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 1**

Illustration of the effect of different aggregation methods for temporal aggregation (e.g. several samplings within one year). Risk quotients (RQ) and assessment values (val) for three different examples (rows), which could reflect different substances or different sites, in 10 samples (S1-S10, columns, e.g. consecutive composite samples) and aggregated values over the 10 samples with minimum aggregation (min), additive-minimum (add-min) aggregation with different parameters ($\alpha$=0.25 means 25% arithmetic mean and 75% minimum, $\alpha$=0.5 means 50% arithmetic mean and 50% minimum), geometric-offset aggregation, and additive aggregation/arithmetic mean with equal weights (add) (Langhans et al., 2014; Schuwirth et al., 2018). See Appendix A1 for the equations of aggregation functions. Colors refer to the five water quality classes from bad (red) to high (blue), according to Fig. 2.

| | Individual Samples | | | | | | | | | | Aggregated values | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 | min | add-min $\alpha$=0.25 | add-min $\alpha$=0.5 | geo-off $\delta$=0.01 | add |
| RQ ex. 1 | 0 | 0 | 0.05 | 0.001 | 0.02 | 15 | 0.05 | 0 | 0 | 0 | | | | | |
| RQ ex. 2 | 0 | 0 | 0.05 | 0.2 | 4.5 | 15 | 7.5 | 1.1 | 0 | 0 | | | | | |
| RQ ex. 3 | 9.5 | 11.5 | 13 | 8.5 | 12.5 | 15 | 14.5 | 7 | 5 | 11.5 | | | | | |
| val ex. 1 | 1.00 | 1.00 | 0.90 | 1.00 | 0.96 | 0.10 | 0.90 | 1.00 | 1.00 | 1.00 | 0.10 | 0.30 | 0.49 | 0.78 | 0.89 |
| val ex. 2 | 1.00 | 1.00 | 0.90 | 0.78 | 0.34 | 0.10 | 0.26 | 0.58 | 1.00 | 1.00 | 0.10 | 0.25 | 0.40 | 0.57 | 0.70 |
| val ex. 3 | 0.21 | 0.17 | 0.14 | 0.24 | 0.15 | 0.10 | 0.11 | 0.28 | 0.33 | 0.17 | 0.10 | 0.12 | 0.14 | 0.18 | 0.19 |

**Table 2**

Useful functions to aggregate the values of sub-objectives to a higher-level objective (Schuwirth et al., 2018; Haag et al., 2019). The last column refers to the properties introduced in Section 2.

| function | equation | fulfilment of properties (see text) |
|---|---|---|
| minimum aggregation | $v_{\min} = \min(v)$ <br> with $v = (v_1, \ldots, v_n)$ values of the sub-objectives to be aggregated | 1, 2, 3, 6 (a and b), 7 |
| additive aggregation | weighted arithmetic mean <br> $v_{\text{add}} = \sum\limits_{i=1}^{n} w_i v_i$ <br> with weights, $w_i$, of the sub-objectives $i$ summing up to 1 | 1, 2, 3, (4), 7 |
| additive-minimum aggregation | $v_{\text{add-min}} = \alpha \cdot v_{\text{add}} + (1-\alpha) \cdot v_{\min}$ <br> with $\alpha$ parameter $>0$ and $<1$ that determines the contribution of the additive aggregation to the overall value | 1, 2, (3), 4, 5, 6a, 7 (for $\alpha < 1$) |
| geometric-offset aggregation | $v_{\text{geo-off}} = (\prod\limits_{i=1}^{n} (v_i + \delta)^{w_i}) - \delta$ <br> with $\delta$ parameter $> 0$ that determines how much compensation between sub-objectives is possible: a value of zero leads to the weighted geometric mean (which has the often undesirable property that the aggregated value is 0 as soon as one of the values of the sub-objectives is 0) and a value of infinity leads to the weighted arithmetic mean; with weights $w_i$ summing up to 1 | 1, 2, (3), 4, 5, 6a, 7 (for $\delta > 0$) |
| Weighted power mean (root-mean-power) | $v_{\text{WPM}} = \begin{cases} \left( \sum\limits_{i=1}^{n} w_i v_i^{\gamma} \right) & \text{for } \gamma \neq 0 \\ \prod\limits_{i=1}^{n} v_i^{w_i} & \text{for } \gamma = 0 \end{cases}$ <br> For $\gamma = \infty$ the weighted power mean leads to the maximum aggregation, for $\gamma = -\infty$ to the minimum aggregation, for $\gamma = 0$ to the geometric mean and for $\gamma = 1$ to the additive aggregation; with weights $w_i$ summing up to 1. | 1, 2, 4, 5, 6a, 7 |

objectives improves (worsens) the aggregated value also improves (worsens).

(5) The aggregation technique helps to provide a differentiated picture, i.e. it facilitates the ranking or grading of sites according to their overall water quality, e.g. to identify hotspots of bad water quality and prioritize management actions.

(6) a) If the value of one of the sub-objectives is in the "bad" class, the overall value shall not be in the "good" or "high" class; or even more extreme: b) if one of the sub-objectives is in the "moderate", "poor" or "bad" class, the aggregated value shall not be in the "good" or "high" class.

(7) Small changes in the fulfilment of sub-objectives should lead to only small changes in the aggregated value. This property is also called "continuity" (Pollesch and Dale, 2015).

These desirable properties are partially conflicting, especially the extreme version of property 6 (b) with properties 4 and 5; there exists no aggregation technique that fulfils all of them to the optimal degree. The *minimum aggregation* fulfils property 3 and is the only of the proposed aggregation methods that fulfils the extreme version of property 6 (b), but it does not fulfil the properties 4 and 5. The *additive aggregation* fulfils property 3 and to some degree 4, but less so property 5 and not property 6. The *additive-*
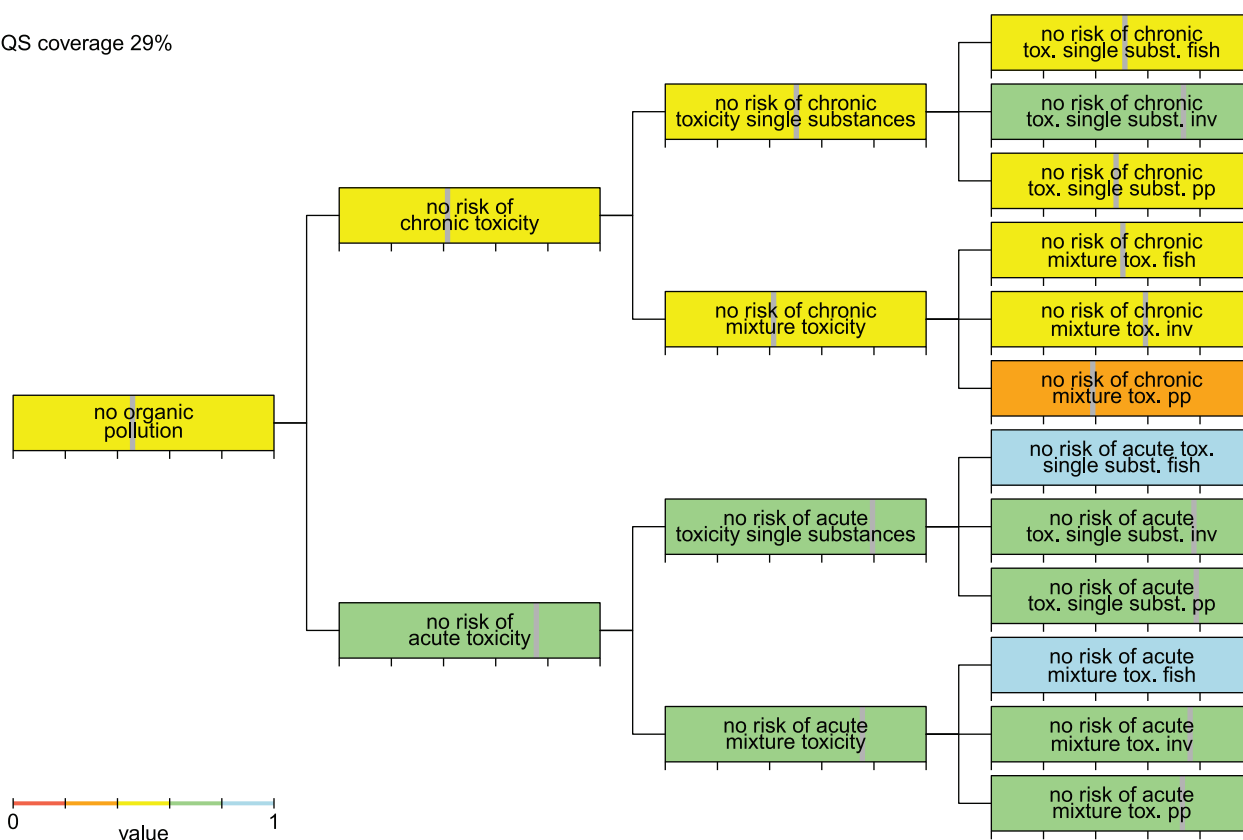
**Fig. 4.** Hierarchical visualization of the assessment results for one sample (site ID 1373, biweekly sampling starting at 11.06.2012) using additive-minimum aggregation with $\alpha = 0.25$ for the aggregation nodes of the first (highest) and third level of the objectives hierarchy (and minimum aggregation for second level as explained in Section 2). Colors refer to the five water quality classes from bad (red) to high (blue), according to Fig. 2; the grey vertical lines indicate the assessment values between 0 and 1 (see legend). The lowest level objectives indicate the organism group: fish, invertebrates (inv), and primary producers / plants (pp). The EQS coverage (top left) indicates, for which percentage of the detected substances an EQS value was available. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

*minimum* and *geometric-offset* aggregation techniques fulfil properties 4 and 5. The fulfilment of property 6 can be adapted based on the choice of the parameter values, but they do not fulfil its extreme version (b). To which degree they fulfil property 3 can be judged differently, but to a lesser degree than *minimum* or *additive* aggregation. In contrast to the aggregation methods mentioned in Table 1, the (weighted) *geometric mean* does not fulfil property 7, because as soon as one of the sub-objectives has a value of zero, the aggregated value is zero. Therefore, a small increase in one sub-objective from zero to a small value above zero can lead to a large difference in the aggregated value, especially if the other sub-objectives have a value close to 1. To avoid this, the *geometric-offset* was introduced (Schuwirth et al., 2018), which prevents this behaviour for $\delta$ values above 0.

Since all of these aggregation techniques fulfil the property 1, they lead to very similar results if all sub-objectives have a similar value (see example 3 in Table 1). However, they lead to very different results if some of the sub-objectives have a very high and others a very low value (see example 1 in Table 1).

The importance of the desirable properties can be controversial among stakeholders and depends on the purpose of the assessment (Langhans et al., 2014). In cases, where it is sufficient to know, if the concentration of any of the substances was above a legal threshold at any time (monitoring compliance with regulation), the *minimum* aggregation is adequate. In cases, where a prioritization among sites is important, when temporal changes should be assessed (trend monitoring), or when responses to management actions should be evaluated (impact control), the *minimum*

*aggregation* is not adequate and the *additive-minimum* or *geometric-offset* aggregation technique is preferable (see Tables 1 and A1). In these cases, the properties 4 and 5 are of particular importance. Haag et al. (2019) provide a method to infer the aggregation method and their parameters from stakeholder interviews. In addition, or as an alternative, the importance of the desirable properties can be discussed and the aggregation function chosen accordingly (Reichert et al., 2019). The parameters can then be chosen based on the evaluation of didactical and real examples that illustrate the consequences of different aggregation functions (see Table 1, and Fig. 7).

To account for the fact that the assessment of single substances can only be equal or better than the assessment of mixture toxicity, we propose an aggregation function for the nodes "no risk of chronic toxicity" and "no risk of acute toxicity" that takes the same value as the nodes "no risk of chronic mixture toxicity" and "no risk of acute mixture toxicity", respectively. (This can be implemented by a *minimum aggregation*.) This means that the assessment of single substances does not affect the higher level objectives, but it serves as additional information to evaluate, if the risk is mainly due to single substances or due to a (complex) mixture of substances (Price and Han, 2011).

To compare the results across time or space, it is important that the number and timing of the samplings as well as the number and selection of substances to be analyzed are standardized within the monitoring program. This poses a challenge for practice, since the list of potentially relevant substances is long and varies over time, as do analytical possibilities, while the budget for monitoring

| | no organic pollution | no risk of chronic toxicity | no risk of chronic toxicity single substances | no risk of chronic tox. single subst. fish | no risk of chronic tox. single subst. inv | no risk of chronic tox. single subst. pp | no risk of chronic mixture toxicity | no risk of chronic mixture tox. fish | no risk of chronic mixture tox. inv | no risk of chronic mixture tox. pp | no risk of acute toxicity | no risk of acute toxicity single substances | no risk of acute tox. single subst. fish | no risk of acute tox. single subst. inv | no risk of acute tox. single subst. pp | no risk of acute mixture toxicity | no risk of acute mixture tox. fish | no risk of acute mixture tox. inv | no risk of acute mixture tox. pp |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1373 2012–03–19 | .48 | .44 | .46 | .40 | .79 | .74 | .44 | .39 | .74 | .61 | .76 | .80 | .81 | .99 | .77 | .76 | .81 | .96 | .73 |
| 1373 2012–04–02 | .36 | .31 | .33 | .40 | .29 | .71 | .31 | .39 | .28 | .55 | .67 | .68 | .95 | .64 | .80 | .67 | .95 | .63 | .72 |
| 1373 2012–04–16 | .43 | .39 | .49 | .52 | .48 | .52 | .39 | .50 | .38 | .40 | .71 | .76 | .97 | .74 | .76 | .71 | .97 | .69 | .71 |
| 1373 2012–04–30 | .41 | .38 | .49 | .54 | .48 | .65 | .38 | .53 | .36 | .40 | .60 | .65 | .95 | .61 | .79 | .60 | .95 | .56 | .71 |
| 1373 2012–05–14 | .35 | .31 | .40 | .45 | .50 | .39 | .31 | .43 | .39 | .29 | .62 | .66 | .98 | .62 | .76 | .62 | .98 | .59 | .59 |
| 1373 2012–05–28 | .29 | .25 | .37 | .40 | .50 | .35 | .25 | .40 | .37 | .22 | .54 | .73 | .97 | .72 | .71 | .54 | .97 | .65 | .48 |
| 1373 2012–06–11 | .46 | .42 | .50 | .51 | .74 | .48 | .42 | .50 | .59 | .39 | .76 | .79 | .98 | .78 | .79 | .76 | .98 | .76 | .73 |
| 1373 2012–06–25 | .45 | .41 | .52 | .52 | .71 | .50 | .41 | .52 | .55 | .38 | .75 | .78 | .97 | .76 | .79 | .75 | .97 | .73 | .73 |
| 1373 2012–07–09 | .55 | .52 | .54 | .50 | .69 | .77 | .52 | .50 | .64 | .59 | .76 | .77 | .98 | .74 | .87 | .76 | .98 | .74 | .78 |

**Fig. 5.** Tabular visualization of the assessment results for one site (ID 1373) and 9 different sampling dates (rows) for all sub-objectives of the hierarchy with additive-minimum aggregation with $\alpha = 0.25$ for the aggregation nodes of the first (highest) and third level of the objectives hierarchy (and minimum aggregation for the second level as explained in Section 2). Numbers indicate the assessment value on a scale from 0 to 1. Colors refer to the five water quality classes from bad (red) to high (blue), according to Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

is limited and not for all substances environmental quality standards are available (Ecotox Centre, 2020). Therefore, it should be foreseen to critically review the monitoring design in this regard from time to time and to adapt the assessment method when necessary.

**Propagation of uncertainty** is technically straightforward using Monte Carlo simulations and can be implemented with any statistical computing language (e.g. R Development Core Team, 2019). However, it requires an estimate of the uncertainty of all inputs to the assessment procedure by the specification of probability distributions, such as attribute levels (here consisting of measured concentrations and environmental quality standards), and aggregation parameters. We can then draw a random sample from these probability distributions and propagate it through the assessment function to derive the uncertainty of the outcome. To which degree this is worth pursuing depends on the scope of the assessment and may be controversial among users. It is probably more relevant for the assessment of future scenarios of management actions based on modelled water quality variables (Schuwirth et al., 2018) and when developing a new monitoring and assessment strategy than for the assessment of routine monitoring results.

Uncertainty in the assessment is sometimes used as an argument for discrete rather than continuous assessments and to favour only two or three instead of five water quality classes. However, the uncertainty of a classification does not only depend on the uncertainty of the attribute level, but also on the location of its expected value. If the expected value of the attribute is close to a class boundary, the uncertainty about the classification can be high, even when the uncertainty about the attribute

value is relatively small (Fig. 3, grey colored distribution). On the other hand, the uncertainty about the classification can be rather low when the expected value is in the middle of a class, even when the attribute has a rather high uncertainty (Fig. 3, shaded distribution).

It is good scientific practice to round numbers of any measurement to provide only significant digits, taking into account the precision of the measurement. For symmetrically distributed values (i.e. not skewed), usually a mean +/- standard deviation is provided. In the examples of Fig. 3, this would be 5.0 +/- 0.8 and 10.0 +/- 0.1, for the shaded and the grey distribution, respectively. This implies that the last provided digit of the mean value is uncertain. Rounding to less than the number of significant digits introduces a discretization error and reduces the information content. This should also be considered when discretizing the results to quality classes. If we provide only two or three instead of five water quality classes, we lose information (which is important to detect changes), but how much this reduces the uncertainty of the classification depends on whether the estimated values are close to a class boundary or not. To judge the uncertainty of the classification, a provision of the attribute value and its uncertainty (e.g. standard deviation or quantile ranges) would be ideal. From this, we can calculate the assessment at the continuous value scale between 0 and 1 as well as its uncertainty, in addition to the estimated quality class. This then also allows us to assess the uncertainty of the classification and to follow changes within one water quality class. A distinction of five quality classes has proven useful (e.g. for ecological assessments in the European Water Framework Directive and in Switzerland) to synthesize numerical results,

| | no organic pollution | no risk of chronic toxicity | no risk of chronic toxicity single substances | no risk of chronic tox. single subst. fish no | risk of chronic tox. single subst. inv | no risk of chronic tox. single subst. pp | no risk of chronic mixture toxicity | no risk of chronic mixture tox. fish | no risk of chronic mixture tox. inv | no risk of chronic mixture tox. pp | no risk of acute toxicity | no risk of acute toxicity single substances | no risk of acute tox. single subst. fish | no risk of acute tox. single subst. inv | no risk of acute tox. single subst. pp | no risk of acute mixture toxicity | no risk of acute mixture tox. fish | no risk of acute mixture tox. inv | no risk of acute mixture tox. pp |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1373 2012–03–19 | .39 | .39 | .40 | .40 | .79 | .74 | .39 | .39 | .74 | .61 | .73 | .77 | .81 | .99 | .77 | .73 | .81 | .96 | .73 |
| 1373 2012–04–02 | .28 | .28 | .29 | .40 | .29 | .71 | .28 | .39 | .28 | .55 | .63 | .64 | .95 | .64 | .80 | .63 | .95 | .63 | .72 |
| 1373 2012–04–16 | .38 | .38 | .48 | .52 | .48 | .52 | .38 | .50 | .38 | .40 | .69 | .74 | .97 | .74 | .76 | .69 | .97 | .69 | .71 |
| 1373 2012–04–30 | .36 | .36 | .48 | .54 | .48 | .65 | .36 | .53 | .36 | .40 | .56 | .61 | .95 | .61 | .79 | .56 | .95 | .56 | .71 |
| 1373 2012–05–14 | .29 | .29 | .39 | .45 | .50 | .39 | .29 | .43 | .39 | .29 | .59 | .62 | .98 | .62 | .76 | .59 | .98 | .59 | .59 |
| 1373 2012–05–28 | .22 | .22 | .35 | .40 | .50 | .35 | .22 | .40 | .37 | .22 | .48 | .71 | .97 | .72 | .71 | .48 | .97 | .65 | .48 |
| 1373 2012–06–11 | .39 | .39 | .48 | .51 | .74 | .48 | .39 | .50 | .59 | .39 | .73 | .78 | .98 | .78 | .79 | .73 | .98 | .76 | .73 |
| 1373 2012–06–25 | .38 | .38 | .50 | .52 | .71 | .50 | .38 | .52 | .55 | .38 | .73 | .76 | .97 | .76 | .79 | .73 | .97 | .73 | .73 |
| 1373 2012–07–09 | .50 | .50 | .50 | .50 | .69 | .77 | .50 | .50 | .64 | .59 | .74 | .74 | .98 | .74 | .87 | .74 | .98 | .74 | .78 |

**Fig. 6.** Tabular visualization of the assessment results for one site (ID 1373) and 9 different sampling dates (rows) for all sub-objectives of the hierarchy with minimum aggregation for all aggregation nodes. Numbers indicate the assessment value on a scale from 0 to 1. Colors refer to the five water quality classes from bad (red) to high (blue), according to Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

| | no organic pollution | no risk of chronic toxicity | no risk of chronic toxicity single substances | no risk of chronic tox. single subst. fish | no risk of chronic tox. single subst. inv | no risk of chronic tox. single subst. pp | no risk of chronic mixture toxicity | no risk of chronic mixture tox. fish | no risk of chronic mixture tox. inv | no risk of chronic mixture tox. pp | no risk of acute toxicity | no risk of acute toxicity single substances | no risk of acute tox. single subst. fish | no risk of acute tox. single subst. inv | no risk of acute tox. single subst. pp | no risk of acute mixture toxicity | no risk of acute mixture tox. fish | no risk of acute mixture tox. inv | no risk of acute mixture tox. pp |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| min | .39 | .39 | .48 | .51 | .74 | .48 | .39 | .50 | .59 | .39 | .73 | .78 | .98 | .78 | .79 | .73 | .98 | .76 | .73 |
| add–min 0.25 | .46 | .42 | .50 | .51 | .74 | .48 | .42 | .50 | .59 | .39 | .76 | .79 | .98 | .78 | .79 | .76 | .98 | .76 | .73 |
| add–min 0.5 | .53 | .44 | .53 | .51 | .74 | .48 | .44 | .50 | .59 | .39 | .78 | .81 | .98 | .78 | .79 | .78 | .98 | .76 | .73 |
| geo–off 0.01 | .63 | .49 | .57 | .51 | .74 | .48 | .49 | .50 | .59 | .39 | .82 | .84 | .98 | .78 | .79 | .82 | .98 | .76 | .73 |
| add | .66 | .49 | .58 | .51 | .74 | .48 | .49 | .50 | .59 | .39 | .83 | .85 | .98 | .78 | .79 | .83 | .98 | .76 | .73 |

**Fig. 7.** Tabular visualization of the assessment results for one sample (site ID 1373, biweekly sampling starting at 11.06.2012) for all sub-objectives (columns) to compare five different aggregation functions (rows) including minimum aggregation (min), additive-minimum aggregation with $\alpha = 0.25$ (add-min 0.25), additive-minimum aggregation with $\alpha = 0.5$ (add-min 0.5), geometric-offset aggregation with $\delta = 0.01$ (geo-off 0.01) and additive aggregation (add) for the aggregation nodes of the first (highest) and third level of the objectives hierarchy (and minimum aggregation for the nodes "no risk of chronic toxicity" and "no risk of acute toxicity", as explained in the text). Numbers indicate the assessment value on a scale from 0 to 1. Colors refer to the five water quality classes from bad (red) to high (blue), according to Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
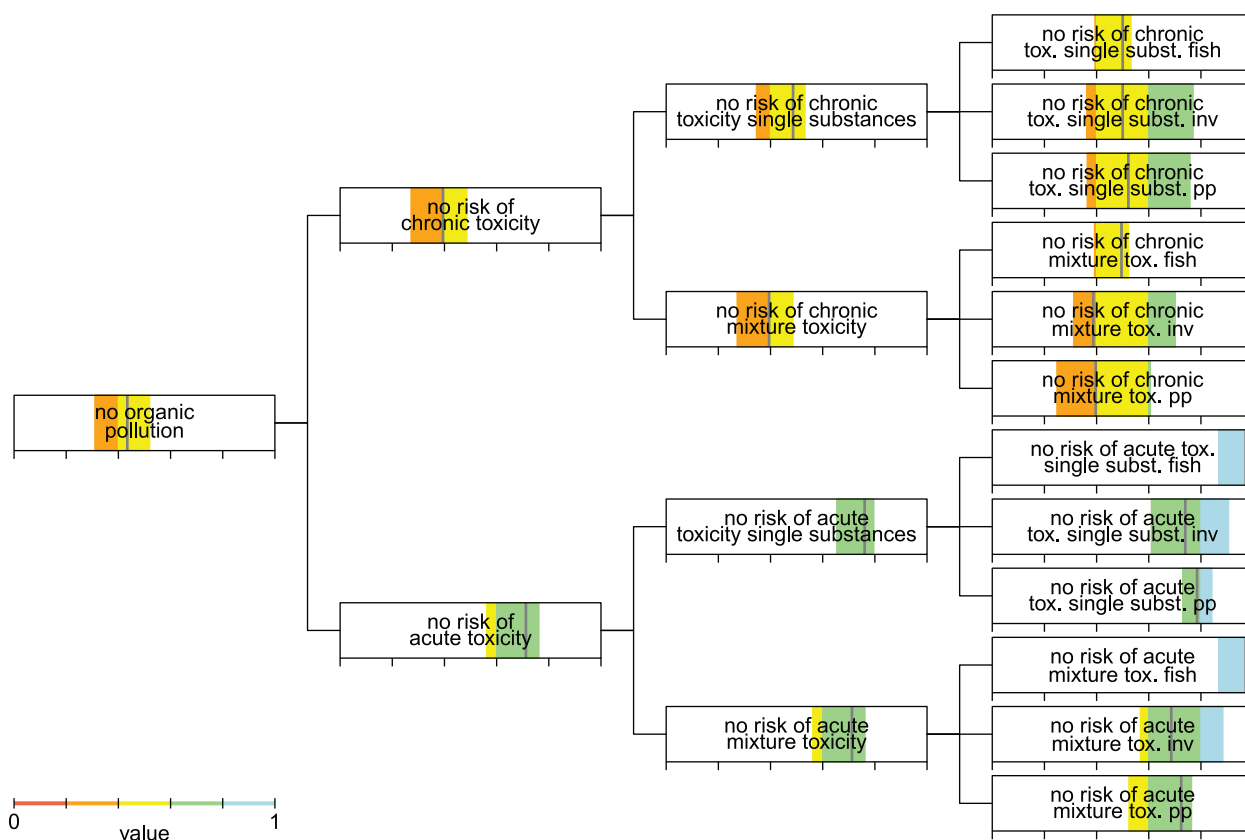
**Fig. 8.** Hierarchical visualization of the assessment results for one site (ID 1373) showing the temporal variability over all nine samples with the colored area for the 5 to 95% range and the median indicated by the grey vertical line (see legend) using additive-minimum aggregation with $\alpha = 0.25$ for the aggregation nodes of the first (highest) and third level of the objectives hierarchy (and minimum aggregation for second level as explained in the text). Colors refer to the five water quality classes from bad (red) to high (blue), according to Fig. 2. Abbreviations in the lowest level objectives indicate the organism group: fish, invertebrates (inv), and primary producers / plants (pp). The same type of visualization can be used to visualize the results of an uncertainty assessment for a single sample. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

e.g. to communicate with policymakers or the public, and to distinguish and prioritize sites.

As mentioned above, the **communication and visualization of results** has to be adapted to the purpose. Often, water quality monitoring programs fulfil multiple purposes. They should provide a representative spatial and temporal overview, identify impaired sites, and ideally support the identification of management actions for remediation. This requires that the assessment method can provide results at each level of the objectives hierarchy to facilitate a synthesis of the results at a high level as well as providing detailed information at the lowest level. Examples for different visualizations for real monitoring data are provided in the next section.

The water quality assessment can be **implemented** with the R-package *utility* (version 1.4.5, Reichert et al., 2013). The R-scripts to produce the assessment results are provided in the supporting information.

The proposed water quality assessment is illustrated for **real monitoring data** from a Swiss federal monitoring program from the year 2012 (Moschet et al., 2014). In this program, biweekly time-proportional composite samples were taken at five different sites in Switzerland. For illustration in this paper, we use the biweekly samples for the assessment of chronic and acute effects. It should be noted, however, that it would be recommended to use samples with a better temporal resolution (e.g. composite samples from 24-96h) for the assessment of acute effects, because an averaging over longer periods usually leads to a dilution and therefore an underestimation of acute effects (Norman et al., 2020). In this monitoring program, 249 different polar organic compounds were analyzed that are sold as pesticides in Switzerland and applied to agricultural or urban land in addition to 134 transformation products, each of which could be quantified in the low ng/L range. In addition, pharmaceuticals were analyzed. However, chronic and acute environmental quality standards were available for only 56 of the analyzed substances (Ecotox Centre, 2020). This means that the water quality status assessment only provides an upper bound and the actual water quality may be worse.

## 3. Illustration of results with real monitoring data

Assessment results for a single sample can be visualized with a color-coded objectives hierarchy (Fig. 4). In this example, the highest risk exists for primary producers from chronic mixture toxicity, which reaches a poor state (orange colored box).

To compare different samples (e.g. a time-series from the same site) a tabular visualization is more convenient (Figs. 5, 6). For this site, single substances that pose the highest risks and exceed chronic EQS are diclofenac (fish), thiacloprid and diazinon (invertebrates), nicosulfurone and metolachlor (primary producers). Similarly, we could plot a table with the results for multiple sites.

The choice of the aggregation functions and their respective parameters for hierarchical aggregation strongly affect the assessment results of the higher level objectives. This can be seen when comparing Fig. 5 with *additive-minimum aggregation* and Fig. 6 with *minimum aggregation* for the same samples. While the

minimum-aggregation clearly reflects the worst value of the sub-objectives, the *additive-minimum aggregation* shows a better differentiation between samples. In Fig. 7, we can compare four different aggregation functions for one sample. In this example, the *minimum aggregation* leads to a "poor" assessment of the main objective ("no organic pollution") with a value of 0.39. In contrast, the *additive aggregation*, which allows for a full compensation between good and bad values of the sub-objectives, leads to a "good" assessment with a value 0.66. The *additive-minimum* aggregation provides a compromise between both, where the parameter $\alpha$ can be adjusted to specify the desired level of compensation (see Table 2 in the main text).

To visualize the results of an uncertainty assessment or the temporal or spatial variability, we can provide the 5 and 95% range of the assessment results as color-coded boxes in the objectives hierarchy (Fig. 8, similarly in a tabular visualization, not shown).

## 4. Conclusions

Multi-attribute value theory provides the methods to formulate a continuous, multi-attribute water quality assessment without any restrictions regarding its functional form (e.g. aggregation functions or non-linear dependence on attributes). This can facilitate a) the explicit formulation and structuring of objectives, b) the consideration of non-linear relationships between measurable attributes and the water quality status, c) the identification of aggregation rules that best represent the dependence of the overall objective on sub-objectives; d) visualization and communication of assessment results at all levels of the objectives hierarchy and at various temporal and spatial scales.

It should be noted that the assessment results strongly depend on the coverage of substances that are monitored and the availability of environmental quality standards for all substances. It would therefore be advisable to complement the water quality assessment with effect-based tools (biotests) and non-target screening of contaminants. Especially for pesticides, that have exposure patterns with a large temporal and spatial variability, the temporal resolution and spatial coverage are crucial aspects of the monitoring design. Furthermore, critical decisions about the hierarchical structure and aggregation methods have to be taken that have consequences for water quality management and policy making. Since the adequacy of aggregation methods depends on the purpose of the monitoring and assessment, it seems reasonable to apply a different aggregation method for compliance monitoring than for the evaluation of temporal and spatial trends or management support.

For the identification of sites that do not comply with regulation, a simple *minimum aggregation* over samplings and sub-objectives is sufficient. The identification of sources of impairment then does not rely on aggregated values but rather on the full time series of all single substances. Stakeholders that have this monitoring purpose in mind are therefore reluctant in accepting any aggregation method that allows for even partial compensation among sub-objectives or samplings, because they do not want a "dilution" of the results.

However, when the scope of the assessment is to get an integrated view on the severity of water quality problems, to identify temporal changes, or detect a response to management actions, an aggregation method that is sensitive to changes in all sub-objectives is more adequate (e.g. the *minimum-additive* or *geometric-offset* aggregation). The visualization and communication of results has to be adapted to the monitoring purpose accordingly.

**Supporting Information**: A zip file with the implementation of the water quality assessment method in the form of a multi-attribute value function, the data from the illustrative example (section 3), and the visualization of the results.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2020.116330.

## References

Baudrot, V., Charles, S., 2019. Recommendations to address uncertainties in environmental risk assessment using toxicokinetic-toxicodynamic models. Sci. Rep. 9, 11432.

Behmel, S., Damour, M., Ludwig, R., Rodriguez, M.J., 2016. Water quality monitoring strategies — a review and future perspectives. Sci. Total Environ. 571, 1312–1329.

Brack, W., Dulio, V., Aring;Gerstrand, M., Allan, I., Altenburger, R., Brinkmann, M., Bunke, D., Burgess, R.M., Cousins, I., Escher, B.I., Hernandez, F.J., Hewitt, L.M., Hilscherova, K., Hollender, J., Hollert, H., Kase, R., Klauer, B., Lindim, C., Herraez, D.L., Miege, C., Munthe, J., O'toole, S., Posthuma, L., Rudel, H., Schafer, R.B., Sengl, M., Smedes, F., Van De Meent, D., Van Den Brink, P.J., Van Gils, J., Van Wezel, A.P., Vethaak, A.D., Vermeirssen, E., Von Der Ohe, P.C., Vrana, B., 2017. Towards the review of the European Union Water Framework Directive: recommendations for more efficient assessment and management of chemical contamination in European surface water resources. Sci. Total Environ. 576, 720–737.

Drakvik, E., Altenburger, R., Aoki, Y., Backhaus, T., Bahadori, T., Barouki, R., Brack, W., Cronin, M.T.D., Demeneix, B., Hougaard Bennekou, S., Van Klaveren, J., Kneuer, C., Kolossa-Gehring, M., Lebret, E., Posthuma, L., Reiber, L., Rider, C., Rüegg, J., Testa, G., Van Der Burg, B., Van Der Voet, H., Warhurst, A.M., Van De Water, B., Yamazaki, K., Öberg, M., Å, Bergman, 2020. Statement on advancing the assessment of chemical mixtures and their risks for human health and the environment. Environ. Int. 134, 105267.

Dyer, J.S., Sarin, R.K., 1979. Measurable multiattribute value functions. Oper. Res. 27, 810–822.

Ecotox Centre. (2020) https://www.ecotoxcentre.ch/expert-service/quality-standards/proposals-for-acute-and-chronic-quality-standards/ (last accessed 13.02.2020).

Eisenfuehr, F., Weber, M., Langer, T., 2010. Rational Decision Making. Springer, Heidelberg.

European Union, 2000. Directive 2000/60/EC of the European Parliament of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Off. J. Eur. Union L327.

Grabisch, M., Marichal, J.-L., Mesiar, R., 2009. Pap E. Aggregation Functions. Cambridge University Press, Cambridge.

Haag, F., Lienert, J., Schuwirth, N., Reichert, P., 2019. Identifying non-additive multi-attribute value functions based on uncertain indifference statements. Omega 85, 49–67.

Junghans, M., Kunz, P., Werner, I., 2013. Toxizität von Mischungen. Aktuelle, praxisorientierte Ansätze für die Beurteilung von Gewässerproben. English translation: toxicity of mixtures. Current practice-oriented approaches for the assessment of surface water samples. Aqua Gas 93, 54–61.

Keeney, R.L., Raiffa, H., 1976. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Wiley, New York.

Kuemmerlen, M., Reichert, P., Siber, R., Schuwirth, N., 2019. Ecological assessment of river networks: from reach to catchment scale. Sci. Total Environ. 650, 1613–1627.

Langhans, S.D., Lienert, J., Schuwirth, N., Reichert, P., 2013. How to make river assessments comparable: a demonstration for hydromorphology. Ecol. Indic. 32, 264–275.

Langhans, S.D., Reichert, P., Schuwirth, N., 2014. The method matters: a guide for indicator aggregation in ecological assessments. Ecol. Indic. 45, 494–507.

Lienert, J., Koller, M., Konrad, J., Mcardell, C.S., Schuwirth, N., 2011. Multiple-criteria decision analysis reveals high stakeholder preference to remove pharmaceuticals from hospital wastewater. Environ. Sci. Technol. 45, 3848–3857.

Marttunen, M., Haag, F., Belton, V., Mustajoki, J., Lienert, J., 2019. Methods to inform the development of concise objectives hierarchies in multi-criteria decision analysis. Eur. J. Oper. Res. 277, 604–620.

Moschet, C., Wittmer, I., Simovic, J., Junghans, M., Piazzoli, A., Singer, H., Stamm, C., Leu, C., Hollender, J., 2014. How a complete pesticide screening changes the assessment of surface water quality. Environ. Sci. Technol. 48, 5423–5432.

Norman, J.E., Mahler, B.J., Nowell, L.H., Van Metre, P.C., Sandstrom, M.W., Corbin, M.A., Qian, Y., Pankow, J.F., Luo, W., Fitzgerald, N.B., Asher, W.E., Mcwhirter, K.J., 2020. Daily stream samples reveal highly complex pesticide occurrence and potential toxicity to aquatic life. Sci. Total Environ. 715, 136795.

Pollesch, N., Dale, V.H., 2015. Applications of aggregation theory to sustainability assessment. Ecol. Econ. 114, 117–127.

Price, P.S., Han, X., 2011. Maximum cumulative ratio (MCR) as a tool for assessing the value of performing a cumulative risk assessment. Int. J. Environ. Res. Public Health 8, 2212–2225.

R Development Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Reichert, P., Niederberger, K., Rey, P., Helg, U., Haertel-Borer, S., 2019. The need for unconventional value aggregation techniques: experiences from eliciting stakeholder preferences in environmental management. EURO J. Decision Process. 7, 197–219.

Reichert, P., Schuwirth, N., Langhans, S., 2013. Constructing, evaluating and visualizing value and utility functions for decision support. Environ. Model. Software 46, 283–291.

Scholten, L., Schuwirth, N., Reichert, P., Lienert, J., 2015. Tackling uncertainty in multi-criteria decision analysis - an application to water supply infrastructure planning. Eur. J. Oper. Res. 242, 243–260.

Schuwirth, N., Honti, M., Logar, I., Stamm, C., 2018. Multi-criteria decision analysis for integrated water quality assessment and management support. Water Res. X 1, 100010.

Schuwirth, N., Stamm, C., Reichert, P., 2012. Incorporation of uncertainty in decision support to improve water quality.. 6th international congress on environmental modelling and software.

Schwarzenbach, R.P., Escher, B.I., Fenner, K., Hofstetter, T.B., Johnson, C.A., Von Gunten, U., Wehrli, B., 2006. The challenge of micropollutants in aquatic systems. Science 313, 1072.

Spycher, S., Mangold, S., Doppler, T., Junghans, M., Wittmer, I., Stamm, C., Singer, H., 2018. Pesticide risks in small streams—how to get as close as possible to the stress imposed on aquatic organisms. Environ. Sci. Technol. 52, 4526–4535.

Swamee, P.K., Tyagi, A., 2000. Describing water quality with aggregate index. J. Environ. Eng. 126, 451–455.

Swamee, P.K., Tyagi, A., 2007. Improved method for aggregation of water quality subindices. J. Environ. Eng. 133, 220–225.

Topping, C.J., Aldrich, A., Berny, P., 2020. Overhaul environmental risk assessment for pesticides. Science 367, 360.

Wittmer, I., Junghans, M., Singer, H., Stamm, C., 2014. Mikroverunreinigungen – Beurteilungskonzept für organische Spurenstoffe aus diffusen Einträgen. English translation: Micropollutants - Concept for the Assessment of Organic Trace Substances from Diffuse Sources. Studie im Auftrag des BAFU. Eawag, Dübendorf http://tinyurl.com/y52nubyp.