



HAL
open science

Spatial aggregation of indicators in sustainability assessments: Descriptive and normative claims

Sandrine Allain, Gaël Plumecocq, Delphine Burger-Leenhardt

► To cite this version:

Sandrine Allain, Gaël Plumecocq, Delphine Burger-Leenhardt. Spatial aggregation of indicators in sustainability assessments: Descriptive and normative claims. *Land Use Policy*, 2018, 76, pp.577-588. 10.1016/j.landusepol.2018.02.038 . hal-02621725

HAL Id: hal-02621725

<https://hal.inrae.fr/hal-02621725>

Submitted on 26 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Spatial aggregation of indicators in sustainability assessments: descriptive and**

2 **normative claims**

3 **Abstract**

4 Indicators are widely used in sustainability assessments. They serve both a descriptive function
5 (i.e., assessing a situation or effects of potential changes) and a normative function (i.e., allowing
6 the expression of value judgments). These functions are usually considered when identifying and
7 using indicators. However, processes such as formalization, estimation, and customization are
8 needed to produce tangible indicators. These processes and their influence on sustainability
9 assessments are studied less often. We focus on spatial aggregation, a specific type of customization
10 commonly used for landscape-scale and regional assessments. Using a database with 146 indicator
11 profiles for water management, we investigated reasons for spatial aggregation choices, i.e. whether
12 indicators based on spatially-explicit data are aggregated while under development or are provided
13 to users in a disaggregated form. Although the literature assigns a descriptive function to spatial
14 aggregation, our database shows that reasons underlying aggregation choices are more diverse.
15 These reasons include highlighting differences, fitting to the scale of a process, fitting to criteria,
16 recognizing a lack of knowledge, expressing social rationality, contextualizing information, and
17 allowing different interpretations of the same indicator. Some of these reasons reflect the choice to
18 expand or reduce the range of potential uses of an indicator, and therefore the potential for different
19 viewpoints to confront each other. Hence, normative claims combine with descriptive claims when
20 aggregating indicators, and even more so when customizing them. In general, the form of indicators
21 merits more attention in the practice and theory of sustainability assessments.

22 **Highlights**

- 23 • Indicators are viewed as objects to describe and debate a situation.
- 24 • Indicators result from different information processes that are sometimes “hidden”.
- 25 • The process of spatial aggregation is investigated.
- 26 • Spatial aggregation choices provide a degree of leeway in interpreting indicators.
- 27 • Choices illustrate tension between the need for consistency and that for diversity.

30 **1. Introduction**

31 Sustainable management of natural resources requires governance that considers long-term
32 dynamics and the spatial scale of the resource managed and that allows different actors to

33 participate in the decision-making process. Sustainability assessments combine tools that can assist
34 decision-makers in this task (Ness et al., 2007). Most sustainability assessment tools require
35 indicators, which can be used to assess a situation and measure progress towards sustainable
36 development (Pires et al., 2017; Singh et al., 2012; Smeets et al., 1999) or are embedded in
37 prospective methods to assess scenarios involving change or policy options (Leenhardt et al., 2012;
38 Singh et al., 2012).

39 Several guidelines and methods exist to identify relevant sets of indicators for sustainability
40 assessments (Alkan Olsson et al., 2009; Bockstaller et al., 2009; Kurka and Blackwood, 2013; Reed
41 et al., 2006; Valentin and Spangenberg, 2000), and the variety of methods reflects the diversity of
42 contexts in which indicators are used. For instance, within “governance contexts” (Hezri and
43 Dovers, 2006), in which people outside of the political elite participate in the decision-making
44 process, indicators cannot be produced according to a clearly identified audience (Hezri and
45 Dovers, 2006). Considering this, several authors claim that identifying indicators should be a social
46 learning process that involves multiple participants (Bell and Morse, 2004; Fraser et al., 2006;
47 Valentin and Spangenberg, 2000).

48 The large number of studies on indicator identification could suggest that defining the indicator
49 set is the only crucial step of information processing, from which comes results of the sustainability
50 assessment. However, other processes occur between defining the relevant indicator set and
51 evaluating a situation or option. For spatial decision support systems, Uran and Janssen (2003)
52 noted that “output sometimes needs simplification, aggregation, structuring, or another form of
53 processing in order for it to be used in a decision-making process. In some systems this is done
54 automatically, or ‘hidden’, so the user is unaware of the fact that an evaluation step has been made”.
55 This highlights the need to clarify and question the entire chain of processes required to develop
56 indicators (and not only to identify them); this is the core motivation for this article.

57 More specifically, we focus on spatial aggregation, a specific process that occurs after an
58 indicator is identified. Spatial aggregation entails changing fine-resolution data into coarser-
59 resolution data (e.g., the entire landscape or region) to derive “meaningful” information. We
60 distinguish “data” from “information” according to definitions of Pahl-Wostl et al. (2013): “‘Data’
61 are symbols, such as the numbers produced by a temperature-measuring device, whereas
62 ‘information’ places data in relation to some meaning that makes them useful (e.g., impacts of July
63 temperature on the yield of a certain crop).” Spatial aggregation is often considered to serve a
64 descriptive purpose: spatial aggregation choices (e.g., which aggregation pathways, which spatial
65 resolution) can depend on the characteristics of the model used (Faivre et al., 2004; Janssen et al.,
66 2009), on the process the indicator intends to describe (Alkan Olsson et al., 2009), or on the
67 expected assessment scale (Chopin et al., 2017). Spatial aggregation clearly differs from indicator

68 aggregation, which involves condensing several indicators into a smaller number of indicators.
69 Indicator aggregation is known to embed normative values and thus go beyond describing social-
70 ecological processes (Böhringer and Jochem, 2007; Mayer, 2008). For instance, ecological
71 economists have demonstrated that using compensatory or non-compensatory rules to aggregate
72 indicators respectively entails a weak or strong conception of sustainability (Garmendia and
73 Gamboa, 2012; Martinez-Alier et al., 1998). However, since ecological economics is as unfamiliar
74 with spatial decision problems (Allain et al., 2017) as spatial planning is with non-equivalent
75 descriptions of a problem (Ramsey, 2009), spatial aggregation of indicators is not studied as a
76 normative process.

77 We consider spatial aggregation both a descriptive process (i.e., translation of information in a
78 formal system) and a normative process (i.e., in which actors express value judgments). To explore
79 how descriptive and normative aspects are combined when considering a sustainability issue with a
80 spatial dimension, we used a database of indicator profiles built to compare water management
81 options.

82

83 **2. Definitions and theory**

84 **2.1 Defining “indicator”**

85 The term “indicator” has multiple definitions (Abbot and Guijt, 1998), and most authors use a
86 finalist or functionalist definition, in which indicators are defined according to their purpose.
87 Consequently, the nature of indicators remains ambiguous: i.e., whether indicators are ideas,
88 variables, objects, variable values, etc.

89 In quantitative assessments, an indicator set is generally displayed in a table containing the names
90 of variables and some additional attributes, such as definitions, scales, calculation methods, and
91 units. In this sense, indicators are variables, and a situation is assessed according to their values
92 (Chopin et al., 2017). However, from a deliberative perspective, “indicators” include outputs of
93 analysis (“The species is not impacted by gear, as a secondary involuntary catch, in any significant
94 way”), variables (“Gross efficiency of the catch (catch/net P.P.)”) or ideas for analyses
95 (“Length/Frequency analysis of catches”) (examples from Douguet et al., 2010). In addition,
96 Meadows (1998) stated that “Indicators can take many forms. They don’t have to be numbers. They
97 can be signs, symbols, pictures, colors.” Consequently, when indicators are identified in contexts
98 with multiple actors, the term “indicator” represents any type of argument, be it a suggested
99 analysis, a quantitative result, a color or a photograph, that empowers someone to assess or judge a
100 situation.

101 Although the vagueness of the term “indicator” may be useful in deliberative contexts, for the
102 purpose of this article, we consider a narrower definition: *an object with meaningful qualitative or*

103 *quantitative information that facilitates learning about a situation and forming a value judgment*
104 *about it.* This definition assumes the existence of one or many indicator developers, i.e. those who,
105 from a heterogeneous set of suggestions (of varying degrees of development) for potential
106 indicators, create a set of formal objects (the indicators).

107

108 **2.2 Steps in indicator development**

109 Our analysis focuses on indicator development, a stepwise process of coding information that
110 occurs between the processes of identifying potential indicators and using them (Fig. 1). We provide
111 insights into indicator identification and use before describing intermediate processes involved in
112 indicator development.

113 The literature on indicator identification highlights qualities that a “good” sustainability indicator
114 should have (Reed et al., 2006), which are generally linked to the latter’s soundness and ease of use.
115 The literature also contains many frameworks developed to derive indicator sets, such as goal-
116 oriented frameworks (Alkan Olsson et al., 2009), multi-scale and systemic frameworks (Astier et
117 al., 2012; Bossel, 1996), and ecosystem-service frameworks (de Groot et al., 2010). Methods to
118 identify indicators are as varied as the frameworks. They can be divided between top-down (a
119 generic framework transposed or adapted to a local context, e.g. Speelman et al. (2007)) and
120 bottom-up approaches (indicator sets are derived from locally relevant issues, and generic
121 knowledge is used to explore these issues, e.g. Fraser et al. (2006)). Both can involve experts and
122 stakeholders, and can be implemented in a deliberative approach involving “extended peer
123 communities” (Funtowicz and Ravetz, 1990) or in a prescriptive approach. Within deliberative
124 approaches, developing an indicator “profile” (i.e., meta-information related to the scientific
125 validity of an indicator and its relevance to the context in which it will be used) is considered a
126 crucial element of knowledge quality (O’Connor and Spangenberg, 2008; Sluijs et al., 2008).

127 Management science studies of the performativity of indicators and their use as management
128 tools in organizations and society (Desrosières, 1997; Espeland and Sauder, 2016) have inspired a
129 literature on the uses of sustainability indicators. Some authors clearly distinguish indicator uses
130 from misuses (Lyytimäki et al., 2013), suggesting that safeguards are required to avoid
131 inappropriate uses. Other authors build on the principle that the users and uses of sustainability
132 indicators are diverse. For instance, Hezri and Dovers (2006) distinguish five types of indicator uses
133 for policy-making in a governance context: instrumental (to solve problems), conceptual (to
134 increase understanding), tactical (as a “substitute for action”), symbolic (as “ritualistic insurance”)
135 and political (to support a pre-defined position). Because we focus on using indicators for group
136 deliberation rather than effective policy-making, we consider it more appropriate to distinguish a
137 descriptive function from a normative function of indicators (see section 2.3).

138 Indicator development proceeds stepwise, with the potential for several iterations and feedback
139 between steps (Fig. 1). After indicator identification (step 1), “formalization” (step 2) is necessary
140 to help select variables and specify characteristics of future indicators. Indicator profiles result from
141 this formalization, and they may be transferred to similar issues or serve as a source of inspiration
142 for different issues. Next, “estimation” (step 3) produces raw data (e.g., values of the variables
143 selected for the case study that are obtained through expertise, simulation or a data search).
144 “Customization” (step 4) includes all activities that provide meaning and tractability to outputs of
145 the estimation (Leenhardt et al., 2012), such as spatial aggregation, classification (aggregating into
146 classes), creation of archetypes, and shaping (e.g., into a map, a graph). The result of customization
147 is called an “indicator” (see section 2.1). The last step, indicator use (step 5), follows development
148 of indicators. It has another nature: it produces knowledge, defined as information embedded in a
149 context of interpretation (Pahl-Wostl et al., 2013). As emphasized by O’Connor and Spangenberg
150 (2008), knowledge *based on indicators* and knowledge *about indicators* can be derived from
151 indicator use. This knowledge about indicators feeds back into the previous steps of indicator
152 development and can help to identify and develop new indicators and discard or modify other
153 indicators.

154

155 2.3 Assessment and evaluation

156 Indicator use can refer to “assessment” or “evaluation” (Fig. 1). Although they are generally used
157 synonymously, we use the terms to indicate two different activities. Assessment is a descriptive
158 activity in which a user an indicator is compared to other estimated values or references. For
159 instance, the workload on different farms may be measured or estimated, which provides a basis for
160 *assessing* which farms have a workload that *surpasses* the national average. Evaluation is the
161 process of assigning a value judgment to the information, such as whether the workload is
162 *acceptable* or *fair*. In short, one *assesses* when stating that the workload is “high” but *evaluates*
163 when claiming that it is “too high”. In contexts with multiple stakeholders, assessment allows for
164 confrontation of different descriptions of a situation, while evaluation allows for confrontation of
165 different value systems.

166 Because confrontations are essential to democratize knowledge-production and decision-making
167 processes (Funtowicz and Ravetz, 1993), we consider that a “good” indicator should be useful both
168 for assessing and evaluating a situation or alternative option. Deliberative sustainability assessment
169 (Frame and O’Connor, 2011) and systemic sustainability analysis (Bell and Morse, 2004) are
170 frameworks for identifying indicator sets that explicitly recognize this duality. However, these
171 frameworks do not mention that the duality of indicators influences both indicator identification and
172 all of the steps that transform indicators into usable forms (Fig. 1). More specifically, because

173 customization generates meaning, it influences indicator use and in return is influenced by the
174 expected uses. Using the specific case of spatial aggregation, our analysis illustrates how normative
175 considerations become relevant in the customization process.

176

177 **2.4 Spatial aggregation**

178 In defining spatial aggregation as the conversion of fine-resolution data into coarser-resolution
179 data (e.g., the entire landscape or region) to derive “meaningful” information, we emphasize
180 changes in scale and in meaning. Different methods can be used to aggregate spatial data (Chopin et
181 al., 2017; Ewert et al., 2011), but discussing them in detail lies outside the scope of this article.
182 Sustainability assessments at regional or landscape scales usually include spatial aggregation, since
183 indicator developers tend to have wide access to spatially-explicit data at these scales.
184 Consequently, the issue of spatial heterogeneity arises (i.e., should they use average values, quantify
185 heterogeneity, or keep the data at the finest resolution?).

186 Systematic aggregation of spatial values into a single value is criticized as an oversimplification
187 of the processes targeted (Scholes et al., 2013) or because of its opacity to decision-makers (Janssen
188 et al., 2005; Uran and Janssen, 2003). Nevertheless, maintaining spatially disaggregated values,
189 generally by displaying them on maps, is costly and introduces a high level of cognitive complexity
190 that can be difficult for users to handle (Jankowski et al., 2001). Similarly, Uran and Janssen (2003)
191 suggested that providing disaggregated data is useful when users have sufficient skill in processing
192 spatial data or when they are guided in their interpretation.

193 However, while we recognize diverse uses and multiple users of indicators, it seems difficult to
194 target the use of (at least some) indicators to a few specific users. Therefore, to enrich the debate
195 over spatial aggregation, we assume that descriptive and normative uses cannot be separated and
196 that users cannot be defined beforehand.

197

198 **3. Materials and Methods**

199 Here we present (i) how we built the database of indicator profiles on which we based our spatial
200 aggregation analysis, (ii) the database itself, and (iii) the objective and method of the analysis.

201

202 **3.1 Origin of the database of indicator profiles**

203 The database of indicator profiles was developed by implementing steps 1 and 2 described above
204 (Fig. 1) to address water management in the downstream area of the Aveyron watershed (southern
205 France). This agricultural area experiences a structural water imbalance in the low-flow period,
206 when low rainfall coincides with maximum water needs for agriculture. Water management is based
207 on water restrictions that are implemented when river flow falls below a specific threshold. Since

208 national and European policies advocate for more structural management options (Erdlenbruch et
209 al., 2013), we created a set of potential indicators to assess contrasting water management options
210 and encourage public debate over future water management in the Aveyron watershed. Because the
211 topic of agricultural water management is complex and prone to conflict in the area, undertaking the
212 sustainability assessment of different scenarios challenges both our ability to represent the social-
213 ecological processes at stake and create the conditions for dialogue between stakeholders.

214 This context drove our choices for developing indicators. Social mistrust and power asymmetries
215 made collective meetings impossible from the outset and a top-down approach risky. Therefore, we
216 opted for a bottom-up approach, based on card-sorting interviews with local and regional
217 stakeholders (Allain et al., 2016), to construct a grid of criteria (Table 1) reflecting the plurality of
218 stakeholder preoccupations. Then, we used these criteria to identify (step 1) and formalize (step 2)
219 the indicator profiles.

220 Although stakeholders sometimes suggested useful indicators or “proto-indicators” (ideas for
221 indicators, with a variable number of guidelines for their scale, representation, calculation, etc.),
222 most indicators were identified through a series of 14 meetings with scientific or technical experts,
223 most of whom worked in the region. We asked the experts to suggest potential indicators that
224 corresponded to the criteria defined by stakeholders, were relevant at the landscape scale, and if
225 possible, could be estimated quantitatively or qualitatively using model simulations, expertise, or by
226 processing geographic information. Expert interviews included questions to facilitate formalization,
227 such as at which spatial and temporal scales the indicator would be relevant.

228 Stakeholder and expert interviews led to identification of 156 proto-indicators. We specified
229 complete profiles for 146 of them based on our own expertise. The remaining 10 reflected
230 arguments stakeholders provided that neither we nor the experts we met with could transform into
231 variables assessable in an *ex ante* manner.

232 Although the same individuals were sometimes interviewed twice, first as a stakeholder to identify
233 criteria and then as a local expert to suggest indicators, we could not push the ideal of
234 democratizing indicator development further. One reason was that local stakeholders, such as
235 reservoir managers or farmers, use many indicators in their everyday life (e.g. the plant-growth
236 stage for farmers to launch irrigation, the level of water in reservoirs for public agents to release
237 water into rivers). Switching from the time and space scale of operational decisions to the scales of
238 the landscape and multiple years is challenging to them. Also, we aimed at assessing scenarios and
239 not real situations. We therefore had to identify and develop indicators not measurable here and
240 now, hence disconnected from what practitioners are accustomed to handle. These barriers can be
241 levered, but the time required to do so was prohibiting (to stakeholders and to us researchers).

242 Hence, for our case study, we, the experts we interviewed, and colleagues who helped specify and
243 interpret indicator profiles served as indicator developers.

244

245

246 **3.2 Database of indicator profiles**

247 The database (see Appendix) contains 146 indicator profiles (123 based on spatially-explicit
248 data), with details on the following attributes:

249 - Indicator definition

250 - Meaning: related criteria, justification for suggesting the indicator

251 - Estimation: estimation method (model simulation, expertise, calculation based on scenario
252 characteristics), scale of raw data (e.g., field, farm, elementary watershed, river, entire
253 landscape)

254 - Customization: type of representation preferred (e.g., simple value, graph, map, illustration,
255 narrative), aggregation scale

256 - Evaluation scale: landscape or landscape + sub-landscape

257

258 **3.2 Analysis of spatial aggregation**

259 We analyzed the indicator profiles in the database with no additional input from stakeholders or
260 experts. We focused on the attributes of 115 “horizontally” aggregated indicators (i.e., those that are
261 based on spatially-explicit raw data provided for multiple locations and thus require aggregation
262 during either indicator customization or use to become relevant at the landscape scale). Indicators
263 referring to only a point location (e.g., a watershed outlet) were not considered horizontally
264 aggregated because a single value at a single point cannot reflect any spatial heterogeneity.
265 Horizontal aggregation differs from “vertical” aggregation, which yields composite indicators or
266 overall sustainability scores (Allain et al., 2017).

267 We analyzed the 115 profiles according to who performs the spatial aggregation: developers
268 (during customization) or users (during use). For the former, developers aggregate raw data to
269 obtain a coarser resolution (partial aggregation) or a landscape-scale indicator (complete
270 aggregation) (Fig. 2). Under partial aggregation, outputs remain heterogeneous. Partial aggregation
271 can remain spatially explicit (e.g., zones) or yield non-spatial classifications (e.g., farm types).
272 Complete aggregation results in a single value for the entire landscape. In user-led aggregation,
273 users aggregate spatially heterogeneous information to assess or evaluate the entire landscape. User-
274 led aggregations are therefore necessarily complete, although some indicators can be used to assess
275 or evaluate landscape zones or classes in addition to the entire landscape.

276 We reflexively clarified reasons underlying developer-led and user-led aggregations, since we
277 were involved in the identification and formalization processes for all indicators, most of the time as
278 indicator developer interviewers but also as developers (see 3.1). Our aggregation choices were not
279 discussed with the stakeholders expected to use them because the research was not part of a joint
280 project in which stakeholders are clearly named. Also, as our experience showed, people get
281 involved and interested at one moment and then, as the research progresses, other people participate
282 in it. This prompted us to consider a wide range of potential users and favor, when possible,
283 indicators prone to diverse interpretations.

284

285 **4. Results**

286 In presenting results for the 115 indicators, we use 15 indicators as examples to illustrate
287 important points¹ (Table 2).

288

289 **4.1 Reasons underlying developer-led spatial aggregation**

290 **4.1.1 Partial spatial aggregation**

291 The database of indicator profiles shows that only six indicators, associated with four criteria
292 (“safety”, “biodiversity”, “adjustment potential”, and “maintaining natural capital”), should result
293 from partial spatial aggregation. For each, raw data estimated at the field or field-islet scale should
294 be aggregated according to the landscape’s hydrological zones (elementary watersheds) or water
295 management zones (e.g., for water restrictions). Reasons underlying partial aggregation were
296 similar to those for aggregation into classes, which indicator developers chose for 24 indicators (for
297 the criteria “local identity”, “wealth and employment”, “long-term adaptability”, “adjustment
298 potential”, and “equity” and “efficiency”). For these indicators, raw data at the field or field-islet
299 scale should be aggregated by crop or farm type.

300 We identified three main reasons for partial aggregation:

- 301 - *to highlight differences* (between zones or classes). For instance, the indicator “area affected
302 by water restrictions at key periods” estimated at the field scale should be aggregated at the

1 1 We often refer to characteristics of “indicators”, but the characteristics are those of profiles and
2 not of the indicators themselves. Indicator profiles (mainly “definition”, “scale of raw data”,
3 “aggregation scale”, “representation”, and “estimation method”) provide guidelines for spatial
4 aggregation that will be usually strictly implemented to produce the indicators, because many
5 calculations are based on long simulations. This constraint, which precludes a trial-and error
6 approach, also explains the time and effort devoted to clarifying aggregation choices. In certain
7 cases, an indicator profile allows for multiple aggregation scales, which can generate several
8 indicators (one per scale). Consequently, although reasons underlying aggregation choices can be
9 analyzed before indicators are produced, quantitative data (the number of indicators concerned)
10 should be considered only indicative.

303 scale of water restriction zones. This would allow users to visualize whether certain zones
304 are more affected than others, and in which zones farmers have less leeway to adjust crop
305 management practices during the year. Likewise, raw data for “irrigation costs per ha”
306 should be aggregated at the crop-type scale to compare crop types and thus provide an
307 indication of how fairly costs are shared.

308 - *to make an indicator more relevant for the targeted criteria.* For instance, the indicator
309 “variation in water stored in soils” estimated at the field scale could be used “as is” to
310 describe the resilience of farms or crops to water stress; however, to produce an indicator
311 informing the “maintaining natural capital” criterion, data should be aggregated at the
312 elementary watershed scale.

313 - *to fit the scale of a process.* For instance, since the indicator “pollution from plant protection
314 practices” describes water pollution caused by agricultural chemicals, raw data should be
315 aggregated at the elementary watershed scale. Aggregation scales could vary to reflect other
316 processes, such as pollution of food products or soils. This reason appears linked to the use
317 of pressure indicators, which are often needed in *ex ante* assessments since direct
318 measurements and model simulations are not always possible.

319 Our study yielded one reason specific to aggregation into classes: to *increase the intelligibility* of
320 indicators. Because the study area has many fields and farms, data at these two scales become
321 unintelligible when considering the entire landscape. Classifying thus renders fine-resolution data
322 tractable.

323

324 **4.1.2 Complete spatial aggregation**

325 The database of indicator profiles contains 79 indicators (among the 11 criteria) that should be
326 completely spatially aggregated. Two reasons identified for partial aggregation also exist for
327 complete aggregation: relevance for the targeted criteria and fit to the process scale. The former is
328 illustrated by the indicator “volume of rainwater returned to the environment”, which qualifies the
329 biophysical system (the water system, in the “maintaining natural capital” criterion). Since it does
330 not seek to indicate the extent to which each water user manages water capital, it should be
331 aggregated at the landscape scale. The second reason (fit to the process scale) is illustrated by the
332 indicator “irrigation capacity of all farms”. Although its raw data, at the farm scale, can be
333 meaningful at this scale, the experts who suggested this indicator referred to the potential to expand
334 irrigation to additional fields and reallocate water within and among farms. It was also meant to
335 address how quickly agriculture can respond to shocks (e.g., a rise in maize price), which justifies
336 aggregation at the landscape scale.

337 We identified three additional reasons for complete aggregation:

- 338 - *easily comparing management options* by providing a single value. For instance, the
339 indicator “impact of water use restrictions on agricultural yields” (“adjustment capacity”
340 criterion) should be aggregated at the landscape scale for two reasons: (1) information
341 would be too complex at smaller scales (e.g., field or farm) and (2) aggregation by farm type
342 or crop type would mask a potential imbalance among the number of each type. Thus, each
343 field’s production should be aggregated to calculate a landscape-scale value. To make this
344 choice, developers assumed a commensuration process, i.e. to consider all crops and all
345 farms equivalent.
- 346 - *addressing uncertainties*. An example related to model and data uncertainties is the indicator
347 “use rate of reservoirs”, based on model predictions. Since data for individual water
348 withdrawals were unavailable, accuracy of the model (MAELIA, Gaudou et al., 2013) could
349 be verified only by comparing its predictions with observed data for river flows and
350 cumulative water withdrawals. Also, allocating a withdrawal point to a field was based on
351 simplified decision rules that are uncertain at the field scale but robust at the landscape
352 scale. Consequently, the indicator could be used only to quantify cumulative use of
353 reservoirs and not use of each reservoir separately. Other uncertainties arise from a lack of
354 knowledge. The indicator “semi-natural elements contributing to water purification” is a
355 good example. It was aggregated at the landscape scale because we simply do not know
356 enough about water purification processes. It is incorrect to assume that the degree of water
357 purification increases linearly as the area of forest or grassland increases. Thus, aggregating
358 this indicator at the landscape scale seems a more accurate approach than not doing so.
- 359 - *prompting expression of social rationality*. The indicator “match between water storage
360 capacity and irrigation needs” (“efficiency” criterion) seeks to reflect how farmers
361 collectively manage the available infrastructure. At a disaggregated scale, it would indicate
362 whether individual farmers use their reservoirs efficiently or not. This is why complete
363 aggregation was preferred.

364

365 **4.2 Reasons underlying user-led aggregation**

366 The 115 horizontally aggregated indicators can be split in two groups after customization:

- 367 - 71 indicators that already exist at the scale of the evaluation and therefore do not require
368 users to aggregate. Most are at the landscape scale (to assess/evaluate the landscape), but
369 some are also at smaller scales and can be used to evaluate/assess both the landscape and
370 sub-landscapes.

371 - 44 indicators require users to aggregate. Indicators in this group are at sub-landscape scales
372 (sometimes also at the landscape scale) but are meant to be used to assess/evaluate only the
373 entire landscape.

374 Several reasons explain why these 44 indicators should not be completely aggregated during
375 customization and why users must perform at least some of the aggregation:

376 - *Visualizing variability or distribution patterns*, allowing users to interpret it/them. This
377 applies mainly to indicators related to the “equity” criterion, such as the indicator “farms
378 concerned with increase/decrease in water costs”. This variability can be expressed spatially
379 (e.g., in maps) but also statistically (e.g., in boxplots) to help users interpret it.

380 - *Promoting expression of diverse (and possibly conflicting) interpretations*. Indicator
381 developers can choose to let users perform the commensurations that the latter judge
382 relevant or to use only some of the information produced. For instance, the indicator “two-
383 year flood flow” (“safety” criterion) should be provided for the outlet of each sub-
384 watershed. Users can consider the indicator value at the most downstream outlet in the
385 watershed to compare water management options “in general” for the entire watershed, or
386 consider the value for the river where they consider safety concerns greater.

387 - *Contextualizing information* provided by the indicator. For instance, the indicator “impacts
388 on the number of recreational water activities” should be provided for each site of activity,
389 even though this diversity could be expressed more simply by the number of potential
390 activities. Presenting the indicator in a spatially-explicit way supports users’ learning
391 capacities (i.e., they can understand better the source of its values) and expression of
392 expertise (i.e., they can appreciate a value better when they know what it refers to). This
393 reason often accompanies the previous reason of diverse interpretations.

394 - *Acknowledging a lack of knowledge*. In this case, the indicator is provided in a spatially
395 disaggregated form, which makes interpretation difficult, nay, impossible, but developers
396 still convey to users that, despite being weak for landscape-scale assessments, the indicator
397 matters. A change in its value reflects a change for a specific location or context, but not
398 enough knowledge is present to infer impact at the landscape scale. For instance, the
399 indicator “nitrate pressure” should be provided at the elementary watershed scale but not for
400 the entire landscape. Nutrient flows are complex processes, which makes it difficult to
401 convert pressure on water quality into a quantified impact. Another example, the indicator
402 “changes in gross margin generated by each type of agricultural production”, considered a
403 proxy for employment and wealth, should be provided for each type of production. This
404 indicator lacked complete spatial aggregation because the influence of gross margin on
405 employment and wealth remains little understood.

406

407 5. Discussion

408 5.1 Many reasons underlie aggregation choices

409 Our analysis empirically illustrates how descriptive and normative claims together influence
410 indicator development and more specifically aggregation choices. We showed that customization,
411 which modifies the information provided by an indicator and its potential uses, is shaped by
412 developers' subjectivities. For instance, by providing users with an indicator in a spatially
413 disaggregated form, developers can decide to widen interpretations of the indicator or can
414 complicate indicator use by making it difficult to interpret. Similarly, by aggregating, developers
415 can limit potential uses of an indicator, for instance by prompting the expression of one type of
416 rationality only, e.g. social rationality over individual interests.

417 Although researchers who study customization of spatial indicators and spatial aggregation
418 problems acknowledge that these processes are non-neutral and affect indicator uses (Janssen et al.,
419 2005; Malczewski and Rinner, 2015; Uran and Janssen, 2003; Walz, 2000), they generally justify or
420 question customization choices in terms of "functionality" (Uran and Janssen, 2003). Accordingly,
421 customization choices result from a trade-off between ease of use and "scientific soundness". These
422 choices could be optimized according to customization rules (e.g., a maximum number of colors, a
423 spatial resolution that allows patterns to be perceived quickly).

424 Among the reasons for spatial aggregation highlighted in our analysis, most of them can be
425 understood in terms of functionality. These reasons include relevance for the criteria, fit to process
426 scale, addressing uncertainty, increased intelligibility, contextualization, increasing the visibility of
427 variability and easily comparing management options. However, the desire to make an indicator
428 more functional does not always completely underlie them. Developers can also intend for users to
429 focus on differences within the landscape by highlighting its heterogeneity or on differences among
430 management options by providing a single quantitative value for each option. Furthermore, other
431 reasons escape the logic of functionality (i.e. increasing indicator's ease of use and ability to
432 describe univocally and accurately a management option for a given criterion), such as promoting
433 expression of diverse interpretations or prompting social rationality. The former reflects a desire to
434 yield multiple indicator uses rather than the only "right" use. The latter reason shows a desire to
435 promote specific indicator uses agreeing with developers' view of the problem.

436 As a result, when indicators are expected to foster deliberation among multiple actors, spatial
437 aggregation cannot be viewed only through the lens of functionality. Functionality primarily makes
438 an indicator accessible to those with different skills in a way that is consistent with the
439 system/problem described. It therefore agrees with a descriptive view of indicators as tools to
440 measure progress towards sustainability. However, when confrontation of value systems is

441 considered another role of indicators (Bell and Morse, 2004; Frame and O'Connor, 2011),
442 customization can influence the degree of leeway available for interpreting indicators.
443 Consequently, customization is one way to increase and/or decrease the number of potential value
444 judgments.

445

446 **5.2 What should drive aggregation choices**

447 By making customization choices, indicator developers influence the size of the “interpretation
448 space” left to users. Developers tend to decrease the space for what they consider indicator
449 “misuses” and to increase the space for what they consider “legitimate uses” (but they can also
450 consider that providing a large amount of space is intrinsically good). The issue then arises of
451 whether certain aggregation choices are more legitimate than others. In other words, what should
452 guide aggregation choices besides functionality? We provide two responses that are related to the
453 collective nature of the problem and the commensuration issue inherent in all information
454 aggregation processes.

455 Environmental appraisal methods can be considered value-articulating institutions (as defined by
456 Vatn (2009)) because they address problems that are collective in nature and that promote
457 expression of different types of rationality. For instance, for managing common goods, methods that
458 promote expression of social rationality rather than the confrontation of individual interests should
459 prevail (e.g., the forum rather than the market) (Vatn, 2009). Fine-resolution, spatially-explicit
460 indicators can encourage expression of individual interests in certain cases since they can, for
461 instance, allow users to consider impacts on an individual’s private property. Customization can be
462 a useful way to switch from individual to social rationality. For instance, partial aggregation could
463 blur individual information or show values in a boxplot to transform spatially-explicit heterogeneity
464 into statistical heterogeneity.

465 However, aggregation, whether spatial or not, implies commensuration, i.e. comparing different
466 objects according to a single measurement unit (Espeland and Stevens, 1998; Martinez-Alier et al.,
467 1998). In spatial aggregation, spatial entities with different values for an indicator are summed or
468 averaged into a single value for the entire area considered. We posit that indicator developers should
469 first address the commensuration issue by considering the extent to which different spatial units are
470 commensurable. It may not be appropriate to make performances of farms on alluvial plains
471 commensurate with those of farms in the hills. Some may advocate that performances can be
472 calculated relative to a location-based potential and that it is acceptable to provide an aggregated
473 value. In contexts with multiple actors, however, the relevance of spatial aggregation as a
474 commensuration process should be framed in normative terms (i.e., whether aggregation helps or
475 hinders confrontation of value judgments). If we consider indicators as “channels for bridging

476 realities and meanings” (Abbot and Guijt, 1998), this indicator would probably be more relevant in
477 a spatially disaggregated form, even though it would make comparisons among management
478 options more complex.

479 Other experiences with the uses of maps (which are by essence spatially-disaggregated
480 representations) in governance contexts support our argument that the commensurability of
481 viewpoints should drive spatial aggregation choices. The use of maps as deliberation support tools
482 (Caron and Cheylan, 2005; Lardon and Piveteau, 2005; Rinner, 2006) show that they can mediate
483 the expression and confrontation of different viewpoints. Similarly, the emergence of participatory
484 or collaborative mapping (Goosen et al., 2007; Jankowski, 2009) strengthens the claim that
485 spatialization is a powerful tool for empowering people, even those who are marginalized. Finally,
486 studies in the field of landscape aesthetics have shown that although landscape metrics are
487 correlated with visual preferences of the landscape, stakeholder groups do not value the same
488 aspects of the landscape (Dramstad et al., 2006; Howley, 2011), which argues in favor of using
489 landscape visualizations when users are not known beforehand. These studies reinforce the idea that
490 spatial representations and other disaggregated forms can help make evaluations more deliberative,
491 under certain conditions. This claim differs greatly from the use of horizontally and vertically
492 aggregated forms (i.e., numbers) in most decision-making tools, including sustainability
493 assessments.

494 When considering the normative role of indicators, aggregation choices can increase the ability of
495 users to interpret indicators according to their values and knowledge while moving beyond
496 individual interests. More generally, although identifying indicator sets is a necessary information
497 compression (Giampietro et al., 2006), indicator customization can help to reintroduce plurality and
498 social rationality into the debate. Finally, there is no contradiction between the normative role and
499 descriptive role of indicators, but there is a tension between the need for consistency (e.g., in the
500 final assessment/evaluation scale(s), in the “common” nature of the problem) and the need to
501 represent diversity (e.g., spatial heterogeneity, plurality of individual values).

502 Therefore, when developing indicators, we recommend developers to investigate the following
503 questions:

- 504 1. What are the scales that are consistent with the processes at stake, the assessment
505 criteria, the level of commensurability between the different spatial units, the nature of
506 the problem tackled and the model used? Answering this “consistency” question bounds
507 the possibilities for aggregation choices. If more than one answer is possible, then other
508 questions arise about the representation of diversity.

- 509 2. Would it enrich the debate or the analysis to visualize heterogeneity or contrasts? (if so,
510 prefer disaggregated forms ; if not, prefer the most aggregated forms, which are
511 generally more easy to handle)
- 512 3. Which forms would help users with the elaboration of their value judgment and make
513 comparisons possible? Between this question and the previous one, a tension might
514 exist. In such cases, developing two indicators, one in an aggregated form and one in a
515 disaggregated form, could allow their concurrent use and stimulate discussions among
516 users and social learning².
- 517 4. If a value judgment appears difficult to access, is the information provided still useful, at
518 least to understand the limitations of the indicator? (if not, the indicator should probably
519 be discarded).

520

521 6. Conclusion

522 Using indicators to assess sustainability within a governance context requires producing them
523 without knowing who will use them and how. This context also entails considering the ability of
524 indicators to foster the expression and confrontation of multiple viewpoints (i.e., their normative
525 quality) and not only their ability to accurately describe situations or processes (i.e., their
526 descriptive quality). Involvement of multiple actors and iterations between the steps of
527 identification, development, and use of indicators is therefore preferred. It is not always possible to
528 perform iterations, however, for instance due to a lack of time, data or knowledge (from developers
529 or potential users) or to the cost of developing certain indicators. In these cases, indicator
530 developers have greater leeway in shaping indicators and therefore the descriptions and judgments
531 the indicators will generate.

532 Indicator development, a stepwise process of coding information, is rarely described in the
533 literature on sustainability assessment. We clarified the processing steps required to develop
534 indicators: formalization, estimation, and customization. We used a case study on water
535 management to investigate spatial aggregation, which is a specific customization process common
536 to most landscape-scale or regional sustainability assessments.

537 Spatial aggregation, the process of changing fine-resolution data into coarser-resolution data to
538 derive meaningful information, partly has a descriptive function. Accordingly, choices are made to
539 increase indicators' ease of use, which generally leads towards aggregated forms (ultimately a

11 2 Although in this article we do not report on the use of indicators but on their development, our experience showed how
12 stimulating it can be to provide stakeholders with two indicators resulting from different aggregation choices. For
13 example, the two indicators presented in Fig. 2, which are based on the same raw data, were used in evaluation workshops
14 with stakeholders. Depending on the indicator used, the participants did not formulate the same value judgments and not
15 with the same facility (it was easier with complete aggregation). Also, the diversity of value judgments among
16 stakeholders was much higher and discussions more intensive with the partially aggregated indicator.

540 single value), and scientific soundness, which may lead towards aggregation or not. However,
541 normative considerations can disrupt aggregation choices. These normative claims argue mainly for
542 disaggregated forms of indicators because aggregation then becomes “user-led”, promoting
543 expression of social incommensurabilities. Indicators can be aggregated, however, to limit
544 expression of users’ individual interests.

545 When descriptive and normative claims are considered together, aggregation choices become
546 more complex and leave more room for developer subjectivity. Completely aggregated forms
547 (single values) do not guarantee that an indicator will be “good for describing” and “good for
548 debating”. Leaving indicators in disaggregated forms, however, is not necessarily the best solution
549 either, even for sustainability problems with a structuring spatial dimension. Spatial aggregation
550 choices illustrate tension between the need for consistency and the need to represent diversity. In
551 this respect, the spectrum of indicator forms (e.g., differing degrees of aggregation, different
552 shapes) merits attention in the practice and theory of sustainability assessments.

553

554 **Acknowledgement**

555 This research is part of a PhD project funded by the French Ministry of Higher Education and
556 Research. We would like to thank the people we interviewed (stakeholders and experts) to build the
557 database of indicator profiles, Michelle and Michael Corson for proofreading this manuscript and
558 two anonymous reviewers for their comments especially helpful to link the theory about indicators
559 and our experience from the case study.

560

561 **References**

- 562 Abbot, J., Guijt, I., 1998. Changing views on change: participatory approaches to monitoring the
563 environment. IIED.
- 564 Alkan Olsson, J., Bockstaller, C., Stapleton, L.M., Ewert, F., Knapen, R., Therond, O., Geniaux, G.,
565 Bellon, S., Correira, T.P., Turpin, N., Bezlepkina, I., 2009. A goal oriented indicator
566 framework to support integrated assessment of new policies for agri-environmental systems.
567 Environ. Sci. Policy, Integrated Assessment of Agricultural and Environmental Policies –
568 concepts and tools 12, 562–572. doi:10.1016/j.envsci.2009.01.012
- 569 Allain, S., Plumecocq, G., Leenhardt, D., 2017. How Do Multi-criteria Assessments Address
570 Landscape-level Problems? A Review of Studies and Practices. Ecol. Econ. 136, 282–295.
571 doi:10.1016/j.ecolecon.2017.02.011
- 572 Allain, S., Plumecocq, G., Leenhardt, D., 2016. La structuration d’une évaluation multicritère pour
573 comparer des scénarios territoriaux de gestion de l’eau. Analyse réflexive sur une démarche
574 de recherche ingénierique. Presented at workshop "Aborder les problèmes d’environnement
575 comme des situations de gestion ?", Strasbourg.
- 576 Astier, M., Garcia-Barrios, L., Galvan-Miyoshi, Y., Gonzalez-Esquivel, C.E., Masera, O.R., 2012.
577 Assessing the Sustainability of Small Farmer Natural Resource Management Systems. A
578 Critical Analysis of the MESMIS Program (1995-2010). Ecol. Soc. 17, 25. doi:10.5751/ES-
579 04910-170325
- 580 Bell, S., Morse, S., 2004. Experiences with sustainability indicators and stakeholder participation: a
581 case study relating to a ‘Blue Plan’ project in Malta. Sustain. Dev. 12, 1–14.
582 doi:10.1002/sd.225
- 583 Bockstaller, C., Guichard, L., Makowski, D., Aveline, A., Girardin, P., Plantureux, S., 2009. Agri-
584 Environmental Indicators to Assess Cropping and Farming Systems: A Review, in:

- 585 Lichtfouse, E., Navarrete, M., Debaeke, P., Véronique, S., Alberola, C. (Eds.), Sustainable
586 Agriculture. Springer Netherlands, pp. 725–738. doi:10.1007/978-90-481-2666-8_44
- 587 Böhringer, C., Jochem, P.E.P., 2007. Measuring the immeasurable — A survey of sustainability
588 indices. *Ecol. Econ.* 63, 1–8. doi:10.1016/j.ecolecon.2007.03.008
- 589 Bossel, H., 1996. Deriving indicators of sustainable development. *Environ. Model. Assess.* 1, 193–
590 218. doi:10.1007/BF01872150
- 591 Caron, P., Cheylan, J.-P., 2005. Donner sens à l’information géographique pour accompagner les
592 projets de territoire : cartes et représentations spatiales comme supports d’itinéraires croisés.
593 *Géocarrefour* 80, 111–122. doi:10.4000/geocarrefour.1031
- 594 Chopin, P., Blazy, J.-M., Guindé, L., Tournebize, R., Doré, T., 2017. A novel approach for assessing
595 the contribution of agricultural systems to the sustainable development of regions with
596 multi-scale indicators: Application to Guadeloupe. *Land Use Policy* 62, 132–142.
597 doi:10.1016/j.landusepol.2016.12.021
- 598 de Groot, R.S., Alkemade, R., Braat, L., Hein, L., Willemsen, L., 2010. Challenges in integrating the
599 concept of ecosystem services and values in landscape planning, management and decision
600 making. *Ecol. Complex., Ecosystem Services – Bridging Ecology, Economy and Social
601 Sciences* 7, 260–272. doi:10.1016/j.ecocom.2009.10.006
- 602 Desrosières, A., 1997. Réfléter ou instituer: l’invention des indicateurs statistiques. *Indic. Socio-
603 Polit. Aujourd’hui Paris L’Harmattan* 15–33.
- 604 Douguet, J.-M., Johnson, P.W., O’Connor, M., Failler, P., Ferraro, G., Chamaret, A., 2010.
605 Evaluating the Social Costs of Fishing Activities in A Deliberative Perspective.
- 606 Dramstad, W.E., Tveit, M.S., Fjellstad, W.J., Fry, G.L.A., 2006. Relationships between visual
607 landscape preferences and map-based indicators of landscape structure. *Landsc. Urban Plan.*
608 78, 465–474. doi:10.1016/j.landurbplan.2005.12.006
- 609 Erdlenbruch, K., Loubier, S., Montginoul, M., Morardet, S., Lefebvre, M., 2013. La gestion du
610 manque d’eau structurel et des sécheresses en France. *Sci. Eaux Territ. Numéro* 11, 78–85.
- 611 Espeland, W.N., Sauder, M., 2016. Engines of Anxiety: Academic Rankings, Reputation, and
612 Accountability. Russell Sage Foundation.
- 613 Espeland, W.N., Stevens, M.L., 1998. Commensuration as a Social Process. *Annu. Rev. Sociol.* 24,
614 313–343.
- 615 Ewert, F., van Ittersum, M.K., Heckeley, T., Therond, O., Bezlepkina, I., Andersen, E., 2011. Scale
616 changes and model linking methods for integrated assessment of agri-environmental
617 systems. *Agric. Ecosyst. Environ., Scaling methods in integrated assessment of agricultural
618 systems* 142, 6–17. doi:10.1016/j.agee.2011.05.016
- 619 Faivre, R., Leenhardt, D., Voltz, M., Benoît, M., Papy, F., Dedieu, G., Wallach, D., 2004.
620 Spatialising crop models. *Agronomie* 24, 205–217. doi:10.1051/agro:2004016
- 621 Frame, B., O’Connor, M., 2011. Integrating valuation and deliberation: the purposes of
622 sustainability assessment. *Environ. Sci. Policy* 14, 1–10. doi:10.1016/j.envsci.2010.10.009
- 623 Fraser, E.D.G., Dougill, A.J., Mabee, W.E., Reed, M., McAlpine, P., 2006. Bottom up and top
624 down: Analysis of participatory processes for sustainability indicator identification as a
625 pathway to community empowerment and sustainable environmental management. *J.
626 Environ. Manage.* 78, 114–127. doi:10.1016/j.jenvman.2005.04.009
- 627 Funtowicz, S.O., Ravetz, J.R., 1993. Science for the post-normal age. *Futures* 25, 739–755.
628 doi:10.1016/0016-3287(93)90022-L
- 629 Funtowicz, S.O., Ravetz, J.R., 1990. Uncertainty and Quality in Science for Policy. Springer
630 Science & Business Media.
- 631 Garmendia, E., Gamboa, G., 2012. Weighting social preferences in participatory multi-criteria
632 evaluations: A case study on sustainable natural resource management. *Ecol. Econ., The
633 Economics of Degrowth* 84, 110–120. doi:10.1016/j.ecolecon.2012.09.004
- 634 Gaudou, B., Sibertin-Blanc, C., Therond, O., Amblard, F., Auda, Y., Arcangeli, J.-P., Balestrat, M.,
635 Charron-Moirez, M.-H., Gondet, E., Hong, Y., others, 2013. The MAELIA multi-agent
636 platform for integrated analysis of interactions between agricultural land-use and low-water

- 637 management strategies, in: International Workshop on Multi-Agent Systems and Agent-
638 Based Simulation. Springer, pp. 85–100.
- 639 Giampietro, M., Mayumi, K., Munda, G., 2006. Integrated assessment and energy analysis: Quality
640 assurance in multi-criteria analysis of sustainability. *Energy*, The Second Biennial
641 International Workshop “Advances in Energy Studies” 31, 59–86.
642 doi:10.1016/j.energy.2005.03.005
- 643 Goosen, H., Janssen, R., Vermaat, J.E., 2007. Decision support for participatory wetland decision-
644 making. *Ecol. Eng.*, Wetland restoration at the Society for Ecological Restoration
645 International Conference in Zaragoza, Spain 30, 187–199.
646 doi:10.1016/j.ecoleng.2006.11.004
- 647 Hezri, A.A., Dovers, S.R., 2006. Sustainability indicators, policy and governance: Issues for
648 ecological economics. *Ecol. Econ.* 60, 86–99. doi:10.1016/j.ecolecon.2005.11.019
- 649 Howley, P., 2011. Landscape aesthetics: Assessing the general publics’ preferences towards rural
650 landscapes. *Ecol. Econ.* 72, 161–169. doi:10.1016/j.ecolecon.2011.09.026
- 651 Jankowski, P., 2009. Towards participatory geographic information systems for community-based
652 environmental decision making. *J. Environ. Manage.*, Collaborative GIS for spatial decision
653 support and visualization 90, 1966–1971. doi:10.1016/j.jenvman.2007.08.028
- 654 Jankowski, P., Andrienko, N., Andrienko, G., 2001. Map-centred exploratory approach to multiple
655 criteria spatial decision making. *Int. J. Geogr. Inf. Sci.* 15, 101–127.
656 doi:10.1080/13658810010005525
- 657 Janssen, R., Goosen, H., Verhoeven, M.L., Verhoeven, J.T.A., Omtzigt, A.Q.A., Maltby, E., 2005.
658 Decision support for integrated wetland management. *Environ. Model. Softw.* 20, 215–229.
659 doi:10.1016/j.envsoft.2003.12.020
- 660 Janssen, S., Ewert, F., Li, H., Athanasiadis, I.N., Wien, J.J.F., Théron, O., Knapen, M.J.R.,
661 Bezlepikina, I., Alkan-Olsson, J., Rizzoli, A.E., Belhouchette, H., Svensson, M., van
662 Ittersum, M.K., 2009. Defining assessment projects and scenarios for policy support: Use of
663 ontology in Integrated Assessment and Modelling. *Environ. Model. Softw.*, Special issue on
664 simulation and modelling in the Asia-Pacific regionSI: ASIMMOD 24, 1491–1500.
665 doi:10.1016/j.envsoft.2009.04.009
- 666 Kurka, T., Blackwood, D., 2013. Participatory selection of sustainability criteria and indicators for
667 bioenergy developments. *Renew. Sustain. Energy Rev.* 24, 92–102.
668 doi:10.1016/j.rser.2013.03.062
- 669 Lardon, S., Piveteau, V., 2005. Méthodologie de diagnostic pour le projet de territoire : une
670 approche par les modèles spatiaux. *Géocarrefour* 80, 75–90. doi:10.4000/geocarrefour.980
- 671 Leenhardt, D., Therond, O., Cordier, M.-O., Gascuel-Oudou, C., Reynaud, A., Durand, P., Bergez,
672 J.-E., Clavel, L., Masson, V., Moreau, P., 2012. A generic framework for scenario exercises
673 using models applied to water-resource management. *Environ. Model. Softw.* 37, 125–133.
674 doi:10.1016/j.envsoft.2012.03.010
- 675 Lyytimäki, J., Tapio, P., Varho, V., Söderman, T., 2013. The use, non-use and misuse of indicators in
676 sustainability assessment and communication. *Int. J. Sustain. Dev. World Ecol.* 20, 385–393.
677 doi:10.1080/13504509.2013.834524
- 678 Malczewski, J., Rinner, C., 2015. *Multicriteria decision analysis in geographic information science.*
679 Springer.
- 680 Martinez-Alier, J., Munda, G., O’Neill, J., 1998. Weak comparability of values as a foundation for
681 ecological economics. *Ecol. Econ.* 26, 277–286. doi:10.1016/S0921-8009(97)00120-1
- 682 Mayer, A.L., 2008. Strengths and weaknesses of common sustainability indices for
683 multidimensional systems. *Environ. Int.* 34, 277–291. doi:10.1016/j.envint.2007.09.004
- 684 Meadows, D.H., 1998. *Indicators and information systems for sustainable development.*
- 685 Ness, B., Urbel-Piirsalu, E., Anderberg, S., Olsson, L., 2007. Categorising tools for sustainability
686 assessment. *Ecol. Econ.* 60, 498–508. doi:10.1016/j.ecolecon.2006.07.023

- 687 O'Connor, M., Spangenberg, J.H., 2008. A methodology for CSR reporting: assuring a
688 representative diversity of indicators across stakeholders, scales, sites and performance
689 issues. *J. Clean. Prod.* 16, 1399–1415.
- 690 Pahl-Wostl, C., Giupponi, C., Richards, K., Binder, C., de Sherbinin, A., Sprinz, D., Toonen, T., van
691 Bers, C., 2013. Transition towards a new global change science: Requirements for
692 methodologies, methods, data and knowledge. *Environ. Sci. Policy, Special Issue:
693 Responding to the Challenges of our Unstable Earth (RESCUE)* 28, 36–47.
694 doi:10.1016/j.envsci.2012.11.009
- 695 Pires, A., Morato, J., Peixoto, H., Botero, V., Zuluaga, L., Figueroa, A., 2017. Sustainability
696 Assessment of indicators for integrated water resources management. *Sci. Total Environ.*
697 578, 139–147. doi:10.1016/j.scitotenv.2016.10.217
- 698 Ramsey, K., 2009. GIS, modeling, and politics: On the tensions of collaborative decision support. *J.*
699 *Environ. Manage., Collaborative GIS for spatial decision support and visualization* 90,
700 1972–1980. doi:10.1016/j.jenvman.2007.08.029
- 701 Reed, M.S., Fraser, E.D.G., Dougill, A.J., 2006. An adaptive learning process for developing and
702 applying sustainability indicators with local communities. *Ecol. Econ.* 59, 406–418.
703 doi:10.1016/j.ecolecon.2005.11.008
- 704 Rinner, C., 2006. Mapping in Collaborative Spatial Decision Making. *Collab. Geogr. Inf. Syst.* 85.
- 705 Scholes, R., Reyers, B., Biggs, R., Spierenburg, M., Duriappah, A., 2013. Multi-scale and cross-
706 scale assessments of social–ecological systems and their ecosystem services. *Curr. Opin.*
707 *Environ. Sustain., Terrestrial systems* 5, 16–25. doi:10.1016/j.cosust.2013.01.004
- 708 Singh, R.K., Murty, H.R., Gupta, S.K., Dikshit, A.K., 2012. An overview of sustainability
709 assessment methodologies. *Ecol. Indic.* 15, 281–299. doi:10.1016/j.ecolind.2011.01.007
- 710 Sluijs, J. van der, Douguet, J.-M., O'Connor, M., Ravetz, J., 2008. Évaluation de la qualité de la
711 connaissance dans une perspective délibérative. *Vertigo - Rev. Électronique En Sci.*
712 *Environ.* doi:10.4000/vertigo.5035
- 713 Smeets, E., Weterings, R., voor Toegepast-Natuurwetenschappelijk, N.C.O., 1999. Environmental
714 indicators: Typology and overview. European Environment Agency Copenhagen.
- 715 Speelman, E.N., López-Ridaura, S., Colomer, N.A., Astier, M., Masera, O.R., 2007. Ten years of
716 sustainability evaluation using the MESMIS framework: Lessons learned from its
717 application in 28 Latin American case studies. *Int. J. Sustain. Dev. World Ecol.* 14, 345–361.
718 doi:10.1080/13504500709469735
- 719 Uran, O., Janssen, R., 2003. Why are spatial decision support systems not used? Some experiences
720 from the Netherlands. *Comput. Environ. Urban Syst.* 27, 511–526. doi:10.1016/S0198-
721 9715(02)00064-9
- 722 Valentin, A., Spangenberg, J.H., 2000. A guide to community sustainability indicators. *Environ.*
723 *Impact Assess. Rev., Assessment Methodologies for Urban Infrastructure* 20, 381–392.
724 doi:10.1016/S0195-9255(00)00049-4
- 725 Vatn, A., 2009. An institutional analysis of methods for environmental appraisal. *Ecol. Econ.* 68,
726 2207–2215. doi:10.1016/j.ecolecon.2009.04.005
- 727 Walz, R., 2000. Development of Environmental Indicator Systems: Experiences from Germany.
728 *Environ. Manage.* 25, 613–623. doi:10.1007/s002670010048

729