Linking deliberative evaluation with integrated assessment and modelling: A methodological framework and its application to agricultural water management

Sandrine Allain, Gaël Plumecocq, Delphine Burger-Leenhardt

Version of Record: https://www.sciencedirect.com/science/article/pii/S0016328720300562
Manuscript_65967ced019cbfd12529e7e4992b11a9

Linking deliberative evaluation with integrated assessment and modelling: a methodological framework and its application to agricultural water management

Sandrine Allain[a], Gaël Plumecocq[b], Delphine Leenhardt[c]

[a]AGIR, Toulouse University, INRA, INPT, INP-EI Purpan, BP 52627, F-31326 Castanet-Tolosan Cedex, France, sandrine.allaingranet@gmail.com

[b]AGIR, Toulouse University, INRA, INPT, INP-EI Purpan, BP 52627, F-31326 Castanet-Tolosan Cedex, France, gael.plumecocq@inra.fr

[c]AGIR, Toulouse University, INRA, INPT, INP-EI Purpan, BP 52627, F-31326 Castanet-Tolosan Cedex, France, delphine.burger-leenhardt@inra.fr

Linking deliberative evaluation with integrated assessment and modelling: a methodological framework and its application to agricultural water management

Abstract:

Dealing jointly with both systemic and social complexity is a challenge that we face with sustainability issues. However, joining these two aspects of complexity is more a "post-normal science" call than an effective practice. Roughly speaking, analytical-systemic approaches focus on the complex-system side, while deliberative approaches deal with the multi-actor side. Our framework for evaluating different alternative actions, called the S²CE, is intended to reconcile the two approaches. It takes as premises: the recognition of plural values, which makes competing actions only weakly comparable; and the existence of irreducible uncertainties, which gives a heuristic status to the future. The framework hybridises an integrated assessment and modelling exercise, with a deliberative multi-criteria evaluation one, and concludes with a reflexive stage. Application to agricultural water management illustrates the strength of this cross-fertilization. First, the use of computer simulations has been found to enrich collective deliberation by helping stakeholders form and transform their value judgments, and by raising critical questions about options that appeared consensual. Second, the deliberative evaluation informed the integrated assessment and modelling by revealing some model limitations, along with the model's unequal capacity to reflect the claims of different stakeholders.

Keywords: *post-normal science and technologies, multi-criteria evaluation, integrated assessment and modelling, agricultural water management, collective deliberation, science-policy interface, transdisciplinarity.*

1. Introduction: clarifying the challenge

"Normal" science and "command and control" management have already attracted half a century of criticism, due to their inability to provide satisfactory answers to sustainability issues. In reaction, post-normal science has offered an appealing science-policy philosophy to approach sustainability-related problems, including GMOs and climate change. Post-normal science is supposed to apply "when facts are uncertain, values in dispute, stakes high and decisions urgent" (Funtowicz & Ravetz, 1990). One of its leading principles is to break the dichotomy between facts ("hard", "unquestionable") and values ("soft", "unscientific") by producing knowledge and evaluating its social robustness within extended peer communities (Funtowicz & Ravetz, 1993). Unravelling and communicating uncertainties is another founding principle of the post-normal science approach (van der Sluijs, 2002).

Post-normal science has found fertile ground for development within diverse fields and for different purposes (Turnpenny, Jones, & Lorenzoni, 2011). Applied to complex systems, it has helped to renew the conception of modelling for decision-making (Allison, Dickson, Fisher, & Thrush, 2018; Haag & Kaupenjohann, 2001). It has also been taken up by social-ecological economists to support an alternative view of sustainability that departs from standard environmental economics (Frame & Brown, 2008; Funtowicz & Ravetz, 1994; Giampietro, Mayumi, & Munda, 2006; Munda, 2004) and promotes deliberative processes for reaching collective choices (Paavola, 2007; Vatn, 2009; Zografos, 2015). Modellers of social-ecological systems and social-ecological economists mainly adopt the idea that sustainability problems are "wicked", but with different standpoints. The former adopt a systemic-analytical approach and put the emphasis on interconnections between problems and between scales (changes at one scale can trigger sometimes drastic changes at other scales, e.g. Tschakert et al., 2014). The latter, favouring deliberative approaches, focus instead on radical uncertainties and the weakly-comparable representations across people (Frame & O'Connor, 2011; Martinez-Alier, Munda, & O'Neill, 1998).

Different standpoints imply different drawbacks, when evaluating different options for change. Modelling and scenario exploration, as typical methods in analytical-systemic approaches, help surfacing feedback loops and the range of potential effects of a change. However the model embodies a shared representation of the social-ecological system under scrutiny (Kok, Rothman, & Patel, 2006; March, Therond, & Leenhardt, 2012; Tschakert et al., 2014; Vervoort, Kok, van Lammeren, & Veldkamp, 2010). This representation is often monolithic and offers little room for the confrontation of different value judgments. On the other hand, in deliberative approaches to sustainability issues, systemic complexity is often largely overlooked. To avoid comparing "apples and oranges", social-ecological economists deal with fragmented and competing representations of the problem at stake. The analysis of the interplay between different issues is too narrowly limited to the consideration of trade-offs between stakeholders' preferences. Overall, systemic-analytical and deliberative approaches should be combined in order to avoid choosing between envisioning counter-intuitive effects of changes and tackling social dilemmas.

It is not a new epiphany that systemic and deliberative approaches should be linked better to meet post-normal challenges (Allain, Plumecocq, & Leenhardt, 2017; Bond, Morrison-Saunders, Gunn, Pope, & Retief, 2015; Rauschmayer & Wittmer, 2006). However, transforming this claim into practice remains problematic. This is what we want to focus on in this article. Before putting forth our argument, we would like to clarify some challenges that, in our opinion, explain the difficulty of linking the two approaches.

First of all, *the role of modelling* appears to be a major bone of contention. Analytical approaches accommodate modelling very well, with for instance the Integrated Assessment and Modelling (IAM) community. Models are seen as tools to foster social learning (Pahl-Wostl et al., 2007) as well as to communicate uncertainties, trade-offs and cross-scale effects (Hamilton, ElSawah, Guillaume, Jakeman, & Pierce, 2015; Pahl-Wostl, Schlumpf, Büssenschütt, Schönborn, & Burse, 2000). Involving stakeholders has become a routine practice, although the stages and degrees of stakeholder involvement vary widely among cases (Voinov et al., 2016). The aim is to assemble different sources of knowledge, either in the form of "soft" models used as intermediary objects or as integrated computer models which connect together different sub-models originally used for other purposes.

However, models (especially computer models) suffer from a long history of "technologies of hubris" (Jasanoff, 2003) and are therefore considered cautiously, if not suspiciously, by social scientists for deliberation purposes. Most of the critics come from the accentuation of framing biases, i.e. when the solution to a problem is already incorporated in its formulation, with model-based indicators (Elgert, 2013), as well as the lack of reflection on ambiguity, in the sense of coexistence of equally-legitimate representations of the same system (Brugnach, Dewulf, Henriksen, & van der Keur, 2011). Petersen et al. (2011) add to it the impracticability of models for assessing innovative or marginal discourses: "Using rigorous methods such as scenario design and system-bounded models only partially allowed for incorporation and processing of the creativity and out-of-the box ideas that were generated by stakeholders." (p. 376).

Second, the *conception of "integration"* is another matter of silent dispute. In many respects, the integration promoted in analytical approaches is incompatible with the treatment of incommensurabilities – the absence of a common metric to compare different issues - emphasized by social-ecological economists (Frame & O'Connor, 2011; Martinez-Alier et al., 1998; Munda, 2004). In IAM, for instance, integration covers different domains, e.g. issues, stakeholders, disciplines, processes and models, and scales (Jakeman & Letcher, 2003) – a list that varies across authors. Integration is generally materialized by a conceptual model, which is a formal representation depicting "our current understanding about the structure and workings of a system" (Argent et al., 2016). To a lesser extent, the same occurs with multicriteria analysis, when conceived out of a deliberative framework. A series of attributes or criteria portrays different aspects of a problem that vary in different directions: the analysis consists of assessing trade-offs and producing a synthetic evaluation. In this case, the multicriteria structure reflects an analytical decomposition of the problem, not the juxtaposition of stakeholders' representations.

However, with integration comes the spectre of a single problem representation, hence mode of valuation, that heterodox economists, and among them social-ecological economists, condemn (Spash, 2009). While integration is also a pivotal topic in ecological economics, it has a specific focus on “problem compressions” (Giampietro, 2003) and the way values are articulated (Martinez-Alier et al., 1998; Vatn, 2009). The different evaluation methods promoted (e.g. social multi-criteria evaluation or deliberative multi-criteria evaluation) have in common their investigation of the social structure of a problem as well as the diversity of values attached to it (De Marchi, Funtowicz, Lo Cascio, & Munda, 2000; Stagl, 2006). In the words of (Frame & Brown, 2008): “the notion of integration is dialogic and multi-perspectival; it is not to be imagined that the dialogue will be one that combines many voices and reduces them to a single, consensual view” (p. 226).

Those distinct perspectives on integration resonate with the way participation is conceived in deliberative versus analytical frameworks. For instance, many IAM scientists consider that one crucial aspect of integration is among different knowledge and viewpoints (Hamilton et al., 2015; van Asselt & Rijkens-Klomp, 2002). To this end, they try to gather experts from different disciplines and local actors to contribute to the development of a conceptual model or of scenarios to assess (Leenhardt et al., 2012; Miller, O’Leary, Graffy, Stechel, & Dirks, 2015; Tschakert et al., 2014). They use different participatory techniques and artefacts to access more remote or less formalized information held by non-scientists (e.g. farmers) and to lift “the barriers to integration” (Tress, Tress, & Fry, 2007). In contrast, in social-ecological economics, deliberation (rather than participation) has a political meaning. It constitutes an *institution* that articulates incommensurable values (Vatn, 2009) and not a *tool* to bring together the different pieces of the same puzzle. Concretely, this means that the artefacts of a deliberative exercise (indicators, scenarios, models, etc.) support the expression and confrontation of value judgments (Allain, Plumecocq, & Leenhardt, 2018; Frame & O’Connor, 2011) and are not just descriptors providing a holistic view of the situation. Hence, integration and participation, which seem to flow naturally from a post-normal science endeavour, do not have the same meaning in analytical and deliberative exercises.

The good news is that there are signs of rapprochement between the different communities; hence start of reconciliation between analytical-systemic and deliberative approaches. In this respect, the work of Salliou et al. (2017) that positions itself as Companion Modelling is a fine illustration. Using Bayesian belief networks to represent perceived causal links between orchard pest management and fruit quality, the modellers chose to have one parametrization of the model per actor. Based on these, they compared the different actors’ views on how different management scenarios could impact the fruit production. Using different models rather than a single integrated one generated dialogue about ambiguity, which was very close to the clarification of incommensurabilities advocated by social-ecological economists.

It would also be a caricature to state that the promoters of deliberation who endorse a radical pluralism perspective, are “anti-modelling”. Frame and Brown (2008) do not close the door to modelling in their review of “post-normal technologies”, especially if it takes place in the framework of explorative futures studies. Information and Communication Technologies are also used to having different stakeholders handle diverse knowledge sources, including numerical models, to elaborate their value judgments (Guimarães Pereira, & Funtowicz, 2006). Recently, Douguet et al. (under review) proposed to use their deliberation support tool in conjunction with a scenario simulation model, but so far with extensive leeway in the articulation of the two tools.

Our introduction has now come full circle. To sum up: both systemic-analytical and deliberative approaches offer valuable insights for the management of wicked problems. Their combination is however challenging because the approaches to modelling and integration differ; some initiatives already show a trend towards rapprochement in order to embrace the part of complexity (systemic or social) that each of them lacks. So, in practice, how do we simultaneously envision potential surprising effects and social dilemmas when comparing different options for change? In this paper, we propose a methodological framework called “S^{2}CE” that explicitly links integrated assessment and modelling of social-ecological systems with deliberative evaluation. Our framework is about how to perform this combination under both the radical pluralism and complex system hypotheses, and how to analyse the results of such an operation. We end with an application in the field of agricultural water management that provides insights for discussing the method.

2. The methodological framework: Evaluation embracing social and systemic complexity (S^{2}CE)

The S^{2}CE reflects the fact that our evaluation (E) framework was designed to embrace both social and systemic aspects (“S^2”) of complex problems (“C”). This framework builds upon the strengths of integrated assessment and modelling, on the one hand, and social/deliberative evaluation, on the other. It acknowledges, as premises:

- the coexistence of different value systems that offer different representations of a problem (value pluralism);

- yet, the possibility to compare competing alternatives in relative terms (“weak comparability”, Munda, 2004);

- the systemic nature of social-ecological problems, which implies that their components are interdependent, provoking either rigidities or surprises when implementing changes (uncertainty);

- yet, the relevance, for present action, of exploring contrasting scenarios of change that are however meaningless for the future (in line with Sardar, 2010, who states that futures studies are "futureless").

S^{2}CE has many proximities with the Integraal framework for sustainability evaluations (O’Connor, Small, & Wedderburn, 2010) , but adds to it specific stages linked to the use of simulation models (Leenhardt et al., 2012), i.e. translation of narratives into simulation inputs and customization of simulation outputs for use by non-experts. The Integraal framework, as a typical example of deliberative evaluation, puts to light convergences and divergences among stakeholder judgments, which is a way of formalizing social complexity. However, non-linear or cascading effects have nearly no chance to integrate stakeholders’ argumentation unless they become legible elsewhere. IAM or other systemic-analytical approaches can generate this legibility. S^{2}CE is therefore an extension of the deliberative evaluation loop with a method focusing on system’s complexity, here IAM. One peculiar feature of S^{2}CE is to give indicators a double status: that of arguments (for the deliberative process) and that of descriptors of changes (for the purpose of systemic analysis).

As a result, S^{2}CE consists of the following 8 stages (Fig. 1):

1. **Problem structuring**: the problem space, theoretically infinite, is turned into a finite number of alternatives, stakeholders and issues. It acquires the structure of a social-choice problem.
2. **Translation for modelling**: alternatives and issues, generally expressed in the form of narratives during the previous stage, are translated into model inputs (e.g. scenarios) and outputs (e.g. indicators to obtain), most of the time quantitative.
3. **Simulation**: although this stage looks like “the computer works”, it actually entails a trial-and-error process, as the people implementing simulations must check the coherence of the outputs, and in many cases launch new simulations with reviewed input files or model code.
4. **Processing and customization**: model outputs are turned into indicators, i.e. an object with meaningful information that facilitates learning about a situation and forming a value judgment about it (Allain, Plumecocq, et al., 2018).
5. **Integrated assessment**: alternatives are assessed using a subset of indicators chosen for their ability to unravel trade-offs, uncertainties and surprising effects. In addition, the underlying mechanisms explaining the differences among indicators and among alternatives are analysed, potentially following complementary simulations.

6. **Multicriteria evaluations in groups**: Knowledge derived from stage 5 is communicated to participants in order to have them understand the potential and limitations of indicators produced through simulations. The evaluation process *stricto sensu* then takes place within homogenous groups of stakeholders. Each group (considered a stakeholder) evaluates the alternatives for each criterion, i.e. they build a multi-criteria matrix. The judgments that constitute the cells of the matrix elaborate on a wide set of indicators, either taken from simulations or added *de novo* by participants.
7. **Aggregation and analysis of the multi-actor multi-criteria matrix**: The multi-actor multi-criteria matrix compiles, along three axes (criteria, alternatives, stakeholders) the multi-criteria matrix of each group. A first necessary step, in a deliberative perspective, is to unravel and discuss the evaluation logic of each group. Aggregation is then required to derive a social evaluation of the alternatives. Different procedures can help this analysis: equity matrixes, outranking etc. But aggregation can be done without mathematical algorithms, by qualifying the general pattern of each alternative e.g. most appreciated, most debated, most contrasted, most consensual etc. and favouring collective deliberation to possibly define the most socially relevant alternative(s).
8. **Crossed analysis**: This final stage aims at cross-fertilizing the results of the integrated assessment and those of the muti-actor multi-criteria evaluation. This cross-fertilization operates in two ways.

   First, from integrated assessment and modelling towards multi-actor multi-criteria evaluation. Some questions that the assessment raises, such as threshold effects or context-dependency, can facilitate reflection on the assumptions upon which a social preference emerged.

   Second, the analysis goes from the multi-criteria multi-actor evaluation towards integrated assessment and modelling. Scrutinizing the use of indicators by stakeholders furthers understanding as to if and how numeric values from simulations are turned into value judgments by different people. Above all, it reveals which stakeholders, which alternatives, and which issues are best supplied by the model. Once those framing biases are made clear, the model can be modified or compared to other modes of expertise (other numeric models, local knowledge, “ground” experience, etc.).

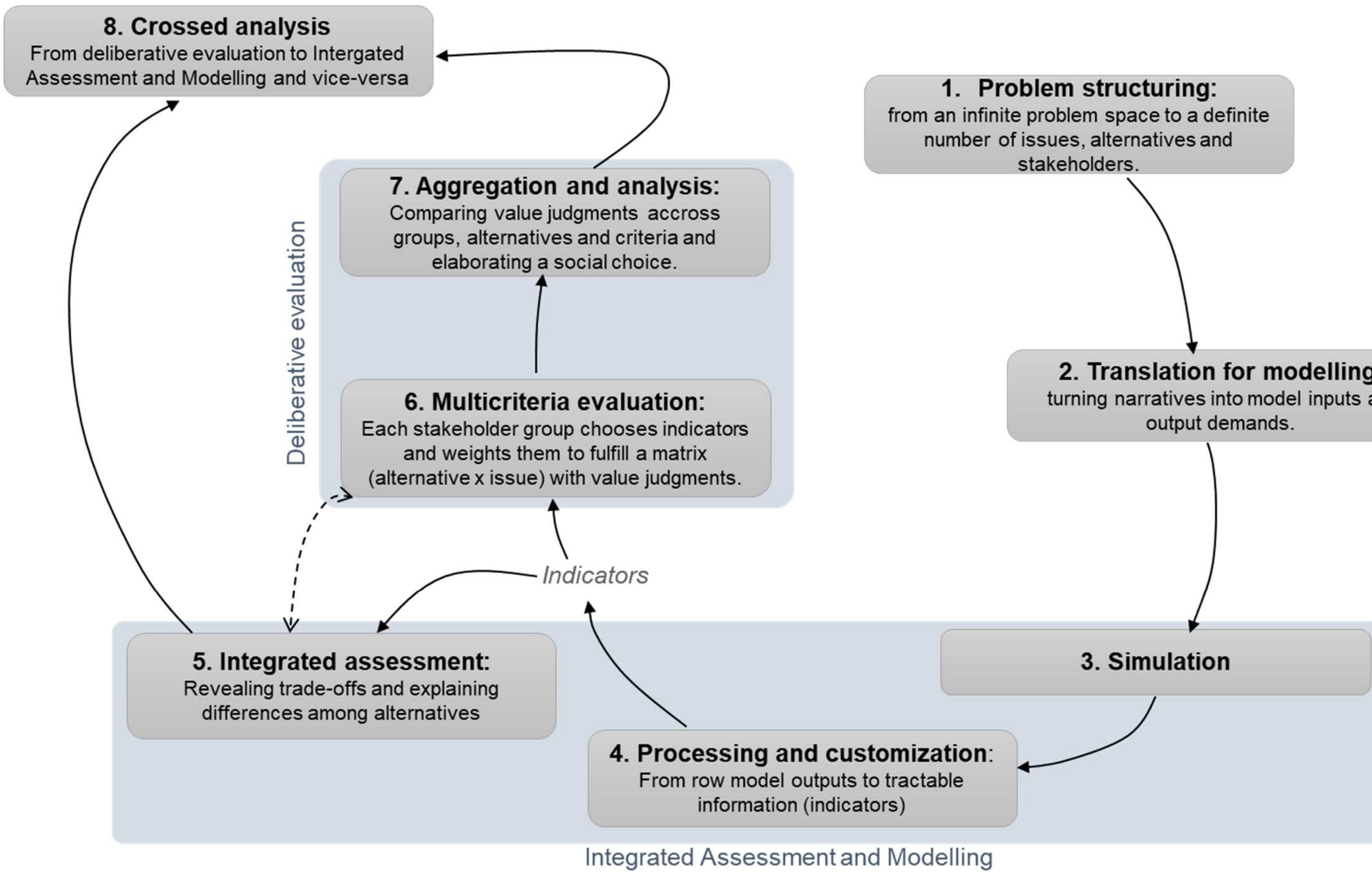


Figure 1. Stages of the S²CE methodological framework

3. Application to agricultural water management

Agricultural water management is an illustrative case of sustainability issues. First, it is one of the topics generating the most intense debates and conflicts throughout the world, even in temperate regions and established democracies (Temper, Delbene, Martinez-Alier, & Rodriguez-Labajos, 2015). Second, water management systems involve numerous and radical uncertainties (Pahl-Wostl et al., 2007) Third, water management systems have encountered multiple failures, whether social, ecological, or both (Budds, 2009; Walker et al., 2002), which make changes all the more urgent.

Calls for a change of paradigm have for more than 20 years sustained the political discourse of an "Integrated Water Resource Management", yet persistent difficulties have been experienced in effectively democratizing water governance and switching the focus from increasing water supply towards regulating water demand. In South Western France, where our case study is situated, there are still strong legacies of the old technocratic paradigm. This is reflected in the prominence of the "Minimum Flow Requirement" indicator (Fernandez, 2014), which currently structures all water management interventions, from the zoning for water policies to the operational decisions to release water stored in public dams or to restrict irrigation withdrawals. It is also evidenced in the persisting idea that water crises in agricultural areas, especially in a context of climatic change, could be solved with more water storage (Debril & Therond, 2012).

With no formalized demand from stakeholders, but building upon long-term collaboration between our research team and water managers, we developed the S²CE framework while reflecting on the specific situation of a "red spot" watershed in our region. This watershed combines severe water imbalance and severe social conflicts. In this section we reconstruct the stages of the S²CE that were implemented there.

3.1. Study area

The Aveyron watershed is one of the Adour-Garonne watersheds exhibiting structural water imbalance, as the Minimum Flow Requirement is regularly not satisfied at the Aveyron outlet. The classification as an imbalanced watershed drives specific policies, among which a legal requirement to reduce water withdrawals. Moreover, every time the river flow falls under the Minimum Flow Requirement, a “drought cell” meets to decide on the level of irrigation restrictions to apply. To avoid or compensate for such crises, flow-supporting reservoirs were built, mostly upstream, in the 1990s and 2000s, and contracts with hydropower companies were signed to ensure the release of specific volumes of water in the rivers during the most critical periods.

Our study focuses more specifically on the downstream portion of the Aveyron watershed (Fig. 2), where most of its irrigated fields are located. Half of the 800 km^2 of the study area is agricultural, and mainly covered by field crops, among which the most demanding in terms of water are maize mono-crops. There are also large patches of fruit trees and seed-maize crops, which provide high added-value but require secured water inputs to be contracted. The maintaining of this agricultural activity is supported both by agricultural companies and by local authorities, as a way to keep this rural territory dynamic and attractive. However, criticism about massive irrigation is also widespread, and a tragic struggle over a reservoir project in the neighbouring watershed contributed to triggering conflict over quantitative water management.

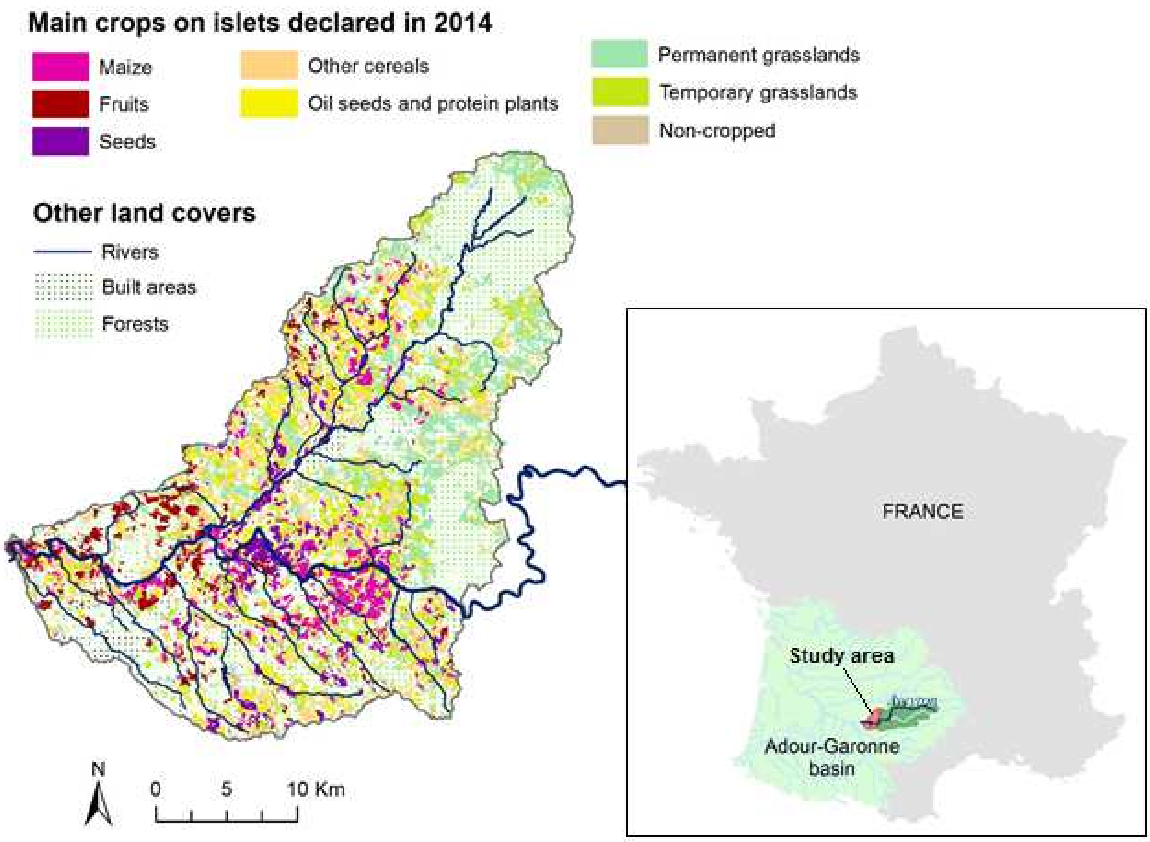


Figure 2. Location of the study area and main land covers

3.2. Method stage 1: Problem structuring

The list of issues and alternatives to consider was built through a bottom-up approach, involving 16 individual interviews based on a card-sorting game (a series of 15 cards with a title and a pictogram, from which interviewees had to choose the 3 to 5 they wanted to express concerns about) and a collective workshop. The 16 stakeholders interviewed were civil servants (from the Water Agency and the regional and local State services in charge of water and the environment), local government agents and councillors (from the district and local communities), environmental associations (including the fishing federation), irrigating farmers, advisors from agricultural extension services and technicians from agricultural cooperatives and suppliers. The choice was based on the principle of "representative diversity" (O'Connor & Spangenberg, 2008), not representativeness.

A collective workshop attended by the interviewees aimed at defining the grid of evaluation criteria (issues) and drawing guidelines for the alternatives to evaluate (Fig. 3). Participants agreed on a list of 11 criteria based on the results of the interviews: safety, food security, economy, biodiversity, local identity, adaptation to exogenous changes, flexibility to adjust the water demand and offer, natural capital, equity, efficiency and political legibility. We also provided them with different levers for quantitative water management and asked them to rank and modify them according to how eager they were to learn about them (independently of their desirability or feasibility). Four alternatives resulting from this workshop were selected: reducing the irrigated area, improving irrigation management at the field scale, generalizing environmentally-friendly cropping systems, and concentrating the water storage capacities.

Figure 3. Workshop with stakeholder to define evaluation criteria (left) and alternatives to evaluate (right)

3.3. Method stage 2: Translation for modelling

In our case, we used the MAELIA multi-agent model (Multi-Agents for EnvironmentaL norms Impact Assessment, Gaudou et al., 2013) which represents the interactions between farming practices, hydrology and water management. The alternatives expressed in the form of narratives by stakeholders had to be translated into input files for the agricultural and hydrological modules of MAELIA. This translation exercise was done "in lab", with no further inputs from stakeholders. The stakeholder alternatives were translated into the following four model-compatible alternatives:

1. Reduction of the irrigated area: irrigated cropping systems were turned into rain-fed cropping systems in the areas not benefitting from flow-support releases. On half of this surface turned to rain-feeding, permanent grasslands were reintroduced.
2. Irrigation using decision-support tools: the decision rules for irrigating field crops were modified in order to follow plant needs and not the actual decision rules of farmers. On the MAELIA platform, it consisted in activating a "theoretical irrigation strategy" which defines the moment for launching irrigation, depending on the soil's humidity and the vegetation stage of the crop. The dose applied was kept unchanged, to be consistent with current irrigation equipment.
3. Crop rotations: Each field with maize mono-cropping (either grain or seeds) was turned into a 4-year rotation alternating sunflower, straw cereals, oil rape and maize.

4. Concentration of water storage capacities: all agricultural reservoirs in the watershed were removed and replaced by three large reservoirs, disconnected from rivers and fed through winter pumping in the Aveyron river (two of the three existing reservoirs were actually enlarged). The total water storage capacity in the watershed remained unchanged, as well as the irrigated surface.

The evaluation criteria had also to be translated into model outputs. This translation occurred in two main stages: the construction of indicator profiles, through expert interviews, and the selection of some for simulation. The indicator profiles (O'Connor & Spangenberg, 2008) comprised, *inter alia*, the names of the indicators, their definition, unit, justification, relevant scales, estimation mode, and representation. Following various exchanges with the modelling team, we selected from the indicator-profile list containing 146 potential indicators (Allain, Plumecocq, et al., 2018) those most susceptible to be impacted by the alternatives, which were easily simulated (with no or nearly no additional model development), using reliable estimates. We ended up with a list of 28 indicators to compute.

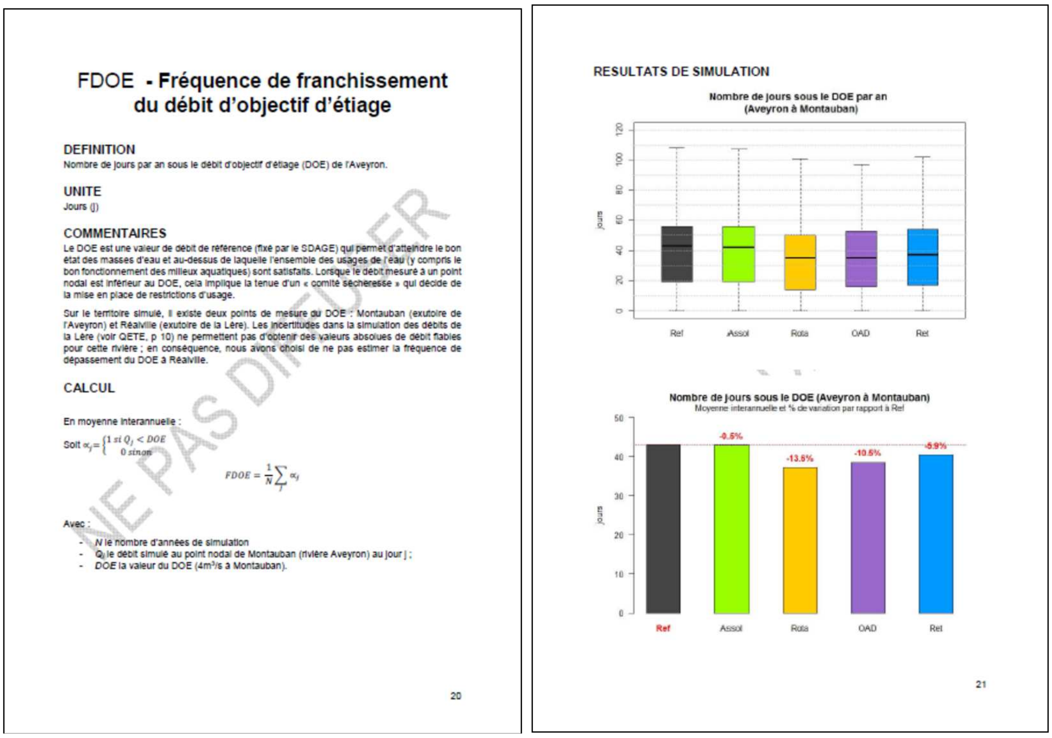
3.4. Method stage 3: Simulations

This stage consisted in running simulations with the MAELIA model for the four alternatives to obtain the outputs necessary to calculate the 28 indicators. Simulation runs were based on a 2001–2013 climate series. If model outputs contradicted the range of values we could expect, we looked for problems either in the input data or in the model code itself. Simulations made us realize for instance how far theoretical water restrictions (calculated by the model according to legal rules) were from real ones (adopted after negotiations). This made us opt for not using this indicator and opting instead for the number of days under Minimum Flow Requirement. Another finding was that the gross margin calculated for maize-seed crops was irrelevant; we therefore discussed our input data with seed breeder trade-unions and modified them accordingly.

3.5. Method stage 4: Processing and customization

From the 28 indicators that we were able to estimate with the model, we created a booklet addressed to the stakeholders[1]. For each indicator, one page detailed the definition of the indicator, its calculation, and its unit of measurement, and commented on the origin of the indicator and its purpose as well as the limitations of the model, especially in terms of reliability. A second page reported the results of the simulation for each water management alternative, in the form of graphs (generally box and bar plots) or maps (Fig.4). In addition to the "indicator booklet", we created an "alternative booklet" summarizing the characteristics of each of the four water management alternatives simulated.

[1] The full booklet (in French) can be found in appendix 7 (p. 285 – 353) of Allain's PhD thesis (Allain, 2018)



FDOE - Fréquence de franchissement du débit d'objectif d'étiage

DEFINITION

Nombre de jours par an sous le débit d'objectif d'étiage (DOE) de l'Aveyron.

UNITE

Jours (j)

COMMENTAIRES

Le DOE est une valeur de débit de référence (fixé par le SDAGE) qui permet d'atteindre le bon état des masses d'eau et au-dessus de laquelle l'ensemble des usages de l'eau (y compris le bon fonctionnement des milieux aquatiques) sont satisfaits. Lorsque le débit mesuré à un point nodal est inférieur au DOE, cela implique la tenue d'un « comité sécheresse » qui décide de la mise en place de restrictions d'usage.

Sur le territoire simulé, il existe deux points de mesure du DOE : Montauban (exutoire de l'Aveyron) et Réalville (exutoire de la Lère). Les incertitudes dans la simulation des débits de la Lère (voir QETE, p 10) ne permettent pas d'obtenir des valeurs absolues de débit fiables pour cette rivière ; en conséquence, nous avons choisi de ne pas estimer la fréquence de dépassement du DOE à Réalville.

CALCUL

En moyenne interannuelle :

Soit $\alpha_j = \begin{cases} 1 \; si \; Q_j < DOE \\ 0 \; sinon \end{cases}$

$$FDOE = \frac{1}{N}\sum_j \alpha_j$$

Avec :

- N le nombre d'années de simulation
- Q_j le débit simulé au point nodal de Montauban (rivière Aveyron) au jour j ;
- DOE la valeur du DOE (4m³/s à Montauban).

20

Figure 4. Example of a double page from the "indicator booklet", here for the indicator "number of days below low-water regulatory flow".

3.6. Method stage 5: Integrated assessment

Our assessment was based on 8 indicators selected from the set of 28, e.g. number of days under the minimum flow requirement at the Aveyron outlet or the total agricultural production for field crops. We chose them specifically to reflect on how changes at the fine scale of fields and water withdrawal points generated impacts (or not) at the macro scale of the watershed. Observing this switch was possible thanks to the multi-agent nature of MAELIA. The subset of indicators was furthermore a way to access interactions between the three sub-models of MAELIA (i.e. water management, hydrology and agriculture), this time using the integrated nature of the model, and to detect any counter-intuitive effects due to these interactions.

We noted for instance a drastic rise of the water consumption with alternative 4 (concentration of water storage). This effect could be explained by two possible causes: either because there was more water in the river, then fewer restrictions, and then more withdrawals from the river; or because of the pooling of the water resources, thus reducing the limitations encountered by farmers using reservoirs. Looking at additional simulation results, we could conclude on the second mechanism. This stage of integrated assessment is further detailed in another article (Allain, Obiang Ndong, Lardy, & Leenhardt, 2018).

3.7. Method stage 6: Multi-criteria evaluation in group

We organized a series of 7 stakeholder workshops with a total of 31 participants:

- Agronomists from a technical institute for field crops
- Agricultural advisors
- Managers of reservoirs supporting agricultural water use (local government and private managers)
- Agents from State services – local level
- Agents from State services – regional level

- Members of associations for environmental protection (local, regional and national organizations)
- Members of the local group of rural communities (project managers and representatives)

The workshop opened with a presentation of the alternatives and some of the simulation results. The group then completed the evaluation matrix for the chosen criteria (Fig. 5) as follows:

1. The participants decided together which indicators they would need, e.g. the most meaningful for the criterion under scrutiny, with a proposed limit of 5 indicators. These indicators could be found in the indicator booklet but as the booklet could not cover all the relevant aspects, they could also be added by the participants.
2. For each cell (alternative x indicator), the participants had to evaluate the performance of the alternative by comparison with the current situation. To do so, they had to choose between five possible judgments depicted by coloured stickers: satisfactory improvement (green), insignificant change (yellow), displeasing deterioration (red), uncertain change or difficult to interpret (blue), do not know (grey).
3. Once the value judgment had been attributed, participants had to allocate a weight to each indicator (for a total of 100%), reflecting the importance of the indicator in the argumentation. Although weight could theoretically vary among alternatives, all stakeholders chose to keep the same weighting system among alternatives.

Figure 5. Evaluation workshop using the indicator booklet (left photo) and evaluation matrix completed by participants with indicators, values judgments (coloured stickers) and weights (right photo).

We (researchers) then informed the online Kerbabel Deliberation Support Tool (Chamaret, O'Connor, & Douguet, 2009) with data from the paper matrixes. In this interface, the social-choice problem is represented as a cube (Fig. 6). Each cell of the cube contains a judgment, represented by a coloured bar. In its variation with indicators, which we used, Kerbabel DST calculates for each cell (stakeholder group x alternative x criterion) a synthetic judgment that corresponds to the majority value judgment across indicators. Hence, the colour of the bar corresponds to the majority judgment (i.e. the colour with the highest cumulative %) and the length of the bar is proportionate to the percentage of this majority judgment. In case two value judgments are equally high, the Kerbabel interface displays the colour of the worst (e.g. red if green and red both weigh 50%).

We sent extracts from the Kerbabel DST matrix to the stakeholder groups in order to have them check their value judgments and discover the synthetic pattern of their evaluation matrix. Stakeholders had the possibility to provide new comments or new judgments by email.

The full multi-actor multi-criteria matrix could not be completed because each group could not evaluate each scenario against each criterion. We ended up with a cube containing many blank cells, but still of interest for analysis because each group started with the criterion of higher relevance to themselves, then the one of second highest relevance and so on, until the end of the workshop.

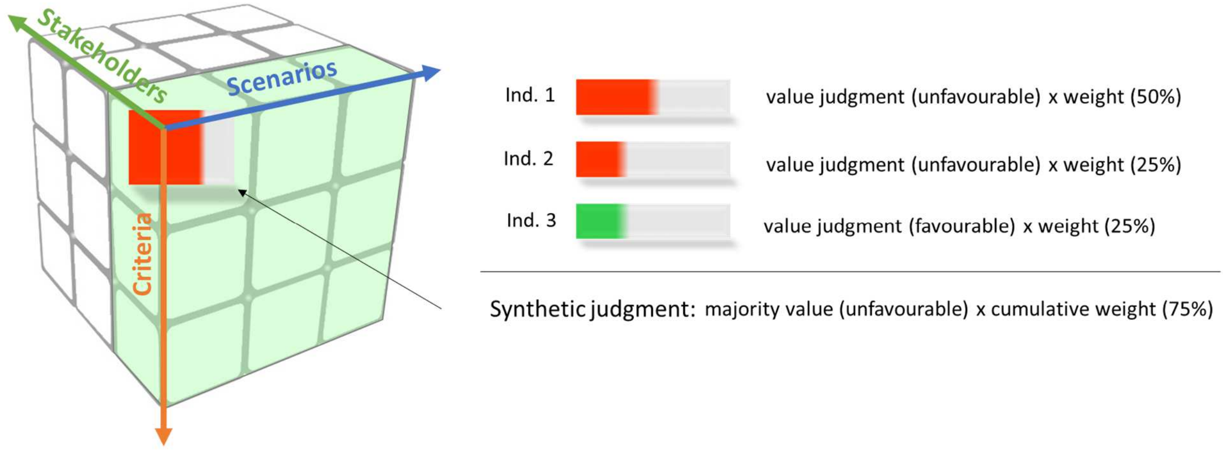


Figure 6. Representation of the multi-actor multi-criteria problem in the Kerbabel DST interface (example of aggregation across 3 indicators).

3.8. Method stage 7: Aggregation and analysis of the multi-actor multi-criteria matrix

3.8.1. In lab

A first aggregation, inside each cell and among indicators, took place while computing synthetic judgments. This aggregation allowed to visualize some general trends in stakeholder groups' preferences (Fig. 7). For instance, for the technical institute, the alternative of decision-support tools appeared to be their favourite, while they liked crop rotations the least. For local communities, by contrast, crop rotations were the alternative they preferred, and the alternative they disliked most was the concentration of water storage. These preferences were further investigated by looking at indicator use. In the case of the economic criterion, agricultural advisors based their judgment mainly on the gross margin of farms and to a lesser extent on other agricultural accounting indicators (e.g. the total revenue of farms, the number of farm employees). Local communities gave a 30% weight to the Aveyron summer flows, in order to highlight the economic importance of non-agricultural activities depending on the Aveyron river (e.g. canoe-kayak, fishing etc.). For the agricultural sector, they used indicators referring to what they considered to be decisive for newcomers to the area (number of farms, revenue generated per m^3 of water withdrawn, etc.). These differences showed different relationships to time (middle-short term vs longer term) and distinct appreciations of the role of agriculture in the local economy (an end vs a means).

For our analysis and the subsequent deliberation, we opted for the entry "by alternative" (rather than "by group" or "by criteria"): based on the synthetic judgments of the different groups, we qualified the general pattern of each alternative. The pattern drawn from Kerbabel DST (Fig. 8) clearly showed two alternatives that were positively distinguished from the others. First, the crop-rotation alternative was the one receiving the highest number of positive judgments (green). Five out of the seven stakeholder groups saw it as a general improvement for all the criteria they evaluated. This alternative was furthermore thought to bring about positive changes overall on 8 of the 10 criteria evaluated. As regards its economic performance, a debate exists as half of the evaluations were positive and the other half negative. The second alternative highlighted by Fig. 8 is the decision-support-tool alternative. This alternative is the only one to be free from negative judgments (i.e. no red cells): there was no major dissent among stakeholder groups nor trade-offs between criteria.

We then created new tables showing the share of each value judgment for each stakeholder group x alternative combination. This way we could visualise the minority judgments that were masked when synthetic judgments are displayed. These tables (Fig. 9) allowed us to see if pros and cons coexisted within one cell. For instance, the alternative "concentration of water storage" had negative judgments (red) disseminated across stakeholder groups and criteria. Even for the group whose preference went to this alternative (agricultural advisers), it bore some weaknesses. The evaluation of the alternative was therefore very sensitive to the weights given to indicators and its social acceptability might be difficult to construct. This alternative appeared moreover as the one bearing the highest percentage of uncertain or unknown value judgments (blue and grey), which means that further knowledge could help clarify the judgments of some groups, and possibly reverse their preferences.

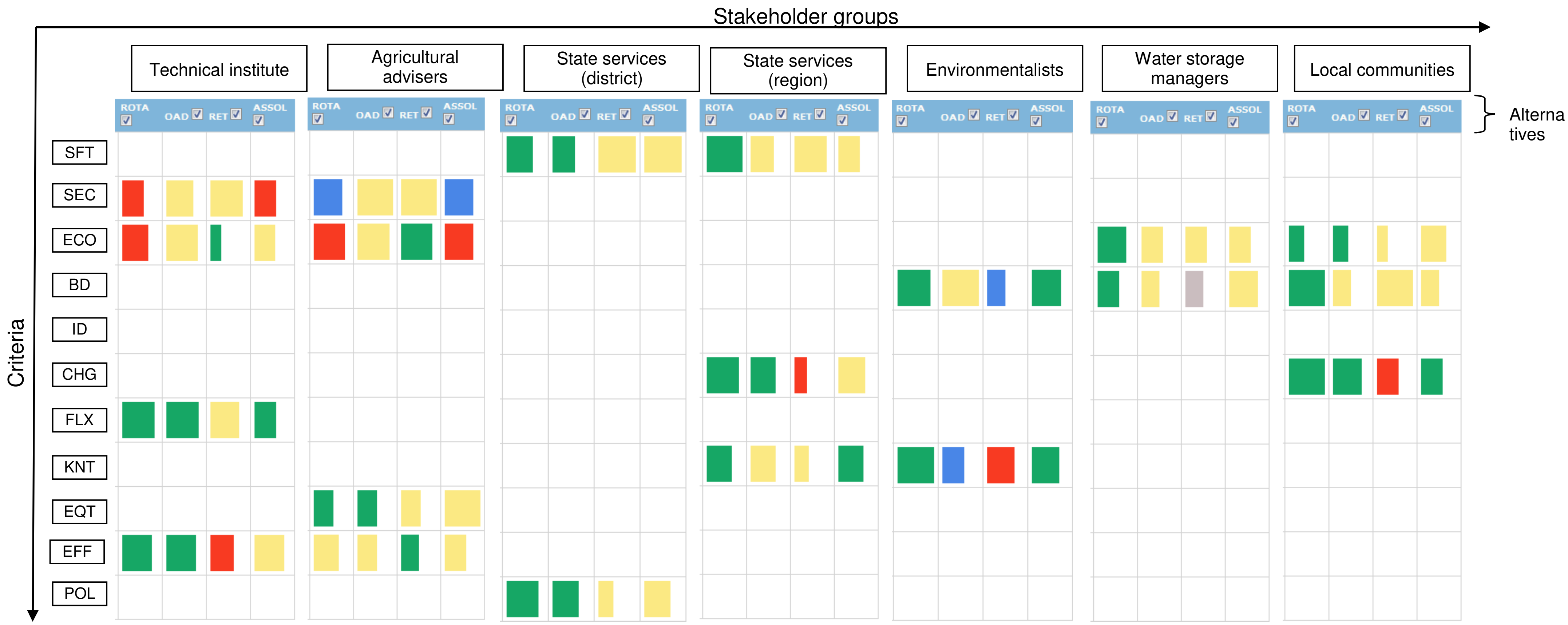


Figure 7. Aggregated results of the evaluation for each stakeholders group (screenshots from the Kerbabel DST web interface). Alternatives under evaluation: crop rotation (ROTA), decision-support tool (OAD), concentration of water storage (RET) and reduction of the irrigated area (ASSOL). Evaluation criteria: safety (SFT), food security (SEC), economy and employment (ECO), biodiversity (BD), local identity (ID – no stakeholder group chose to evaluate as a priority this criterion), adaptation to exogenous changes (CHG), flexibility to adjust the water offer and demand (FLX), natural capital (KNT), equity (EQT), efficiency (EFF), political legibility (POL). The colour code reflects the majority judgment given by a stakeholder group to an alternative for a specific criterion: green = satisfactory improvement; yellow = no significant change; red = displeasing degradation; blue = uncertain; grey = do not know.

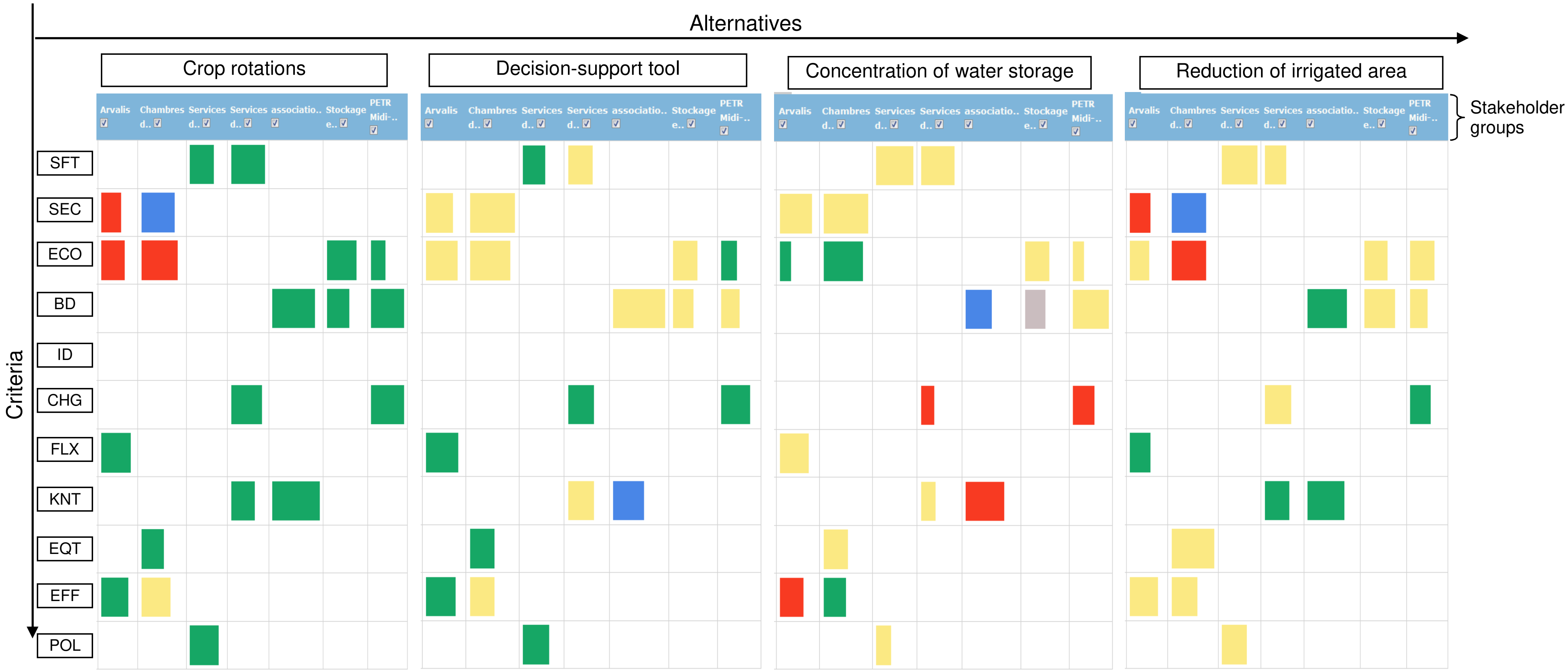


Figure 8. Aggregated results of the evaluation for each alternative (screenshots from the Kerbabel DST web interface). Evaluation criteria: safety (SFT), food security (SEC), economy and employment (ECO), biodiversity (BD), local identity (ID – no stakeholders group chose to evaluate as a priority this criterion), adaptation to exogenous changes (CHG), flexibility to adjust the water offer and demand (FLX), natural capital (KNT), equity (EQT), efficiency (EFF), political legibility (POL). Stakeholder groups (left to right columns): technical institute, agricultural advisers, State services (district), State services (region), environmentalist associations, water storage managers, local communities.

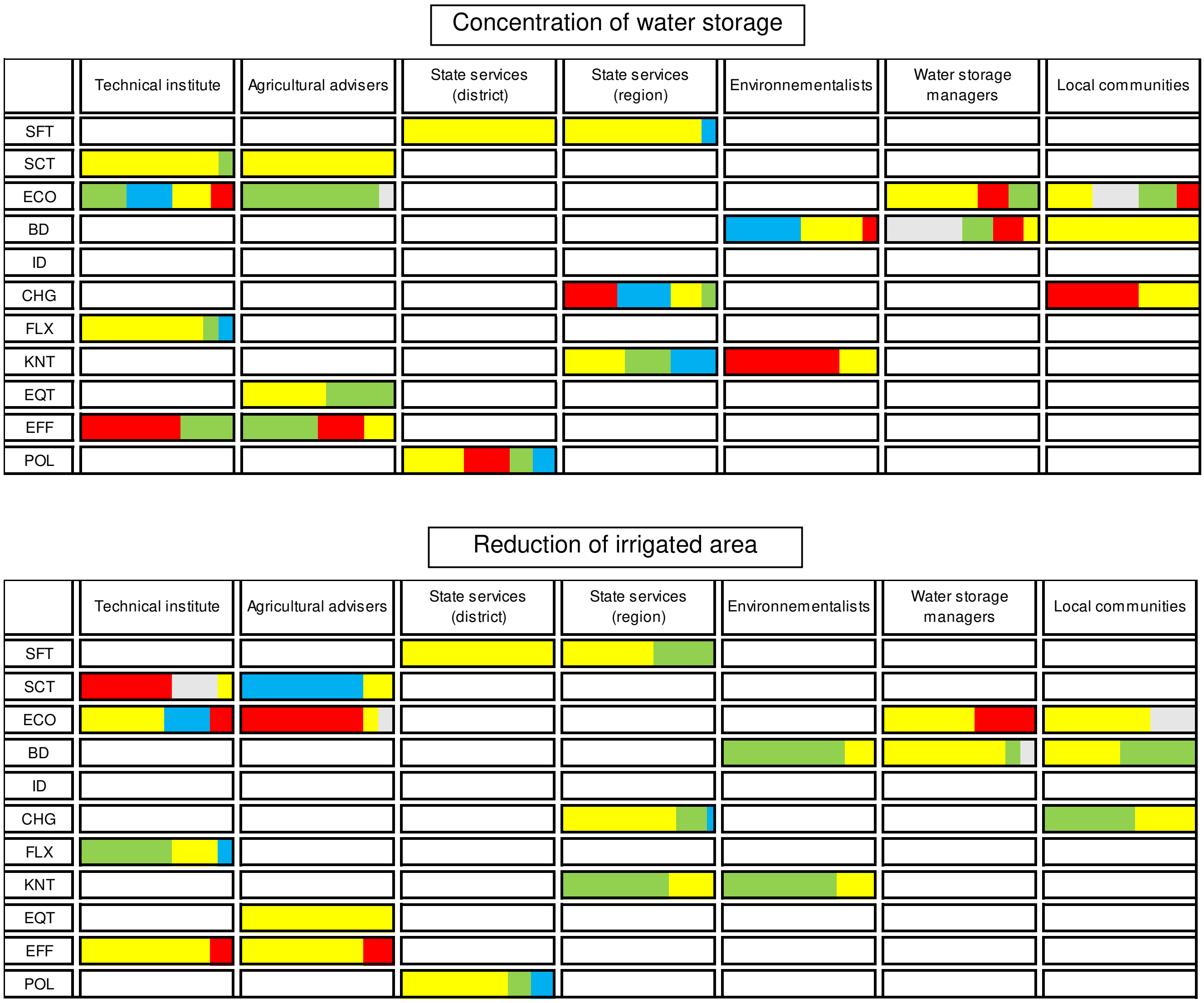


Figure 9. Detail of minority judgments for two alternatives. The colour code reflects the majority judgment given by a stakeholder group to an alternative for a specific criterion: green = satisfactory improvement; yellow = no significant change; red = displeasing degradation; blue = uncertain; grey = do not know.

3.8.2. With stakeholders

The analysis “in lab” was consolidated during the collective restitution that brought together the participants of the workshops. This restitution can be seen as a second step of analysis and aggregation. Our objective was not so much to decide between alternatives, which is generally the final aim of an evaluation exercise, but rather to learn about their social relevance (for whom? for what?).

Participants first discovered the entire cube and the position of their group relative to the other groups. To facilitate this discovery, we displayed results for each alternative on posters, between which participants navigated freely (like at an exhibition). The following plenary discussion focused on the divergences observed, and also the weaknesses and strengths of each alternative. Some participants expressed surprise at the judgments of their peers. This was for instance the case with reservoir managers who were circumspect about the low relevance, for environmentalists, of the alternative “concentration of the water storage capacity”, while it is generally argued that fewer bigger reservoirs improve hydrological functioning. They thus discovered how much importance environmentalists grant to irrigation in itself, as a symbol of agricultural drifts, irrespective of considerations on water flows.

The most salient problems (weak points and points with divergent judgments) revealed by the evaluations were then summed up and two groups were formed to discuss whether these problems could be solved and how. Among the new proposals, one idea was for instance to modify the crop-rotation. The new alternative contained two novelties: (1) replacing maize by alfalfa in the rotation scheme, as alfalfa represents a "productive" grassland that satisfies ecological claims against irrigated maize and agricultural claims for good economic valorisation; and (2) keeping parcels with maize-seed monocrops as a vector for the perpetuation of small farms, an agricultural model that most stakeholders defend, including environmental organizations.

3.9. Method stage 8: Cross analysis

Cross-fertilization between IAM and deliberative evaluation occurred first during the evaluation workshops (stage 6) and analysis of simulations (stage 5) ("hot": in the making), and then more reflexively, after the restitution meeting ("cold", ex post).

3.9.1. "Hot" cross-fertilization

The evaluation workshops provided an opportunity to communicate the most salient simulation results, for instance the drastic increase of water consumption under the "concentration of water reservoirs" alternative. This result astounded State agents and strongly influenced their judgment. Some recognized that the concentration of water reservoirs, which they considered the best option if water management were to be changed from scratch, had turned into a "vanishing dream".

In turn, comments from participants about simulation results raised new questions for the integrated assessment and modelling. For instance, members of the technical institute for field crops expressed surprise about the decrease of water consumption that was simulated under the "decision-support tool" alternative: they said that the simulated decrease was much higher than what they knew from their expertise. Initially this caused us to doubt the credibility of the theoretic irrigation model on which the alternative was built. We realized that it was theoretically correct at the field scale, but that the MAELIA model could not represent "rebound effects" that can occur at a larger scale (farm or landscape scale). Indeed, more efficient irrigation management can lead farmers to increase their irrigated area or to introduce new, more water-demanding crops with higher added-value. When pursuing our integrated assessment work and publishing its results, we therefore insisted on this limitation, which puts into perspective the capacity of decision-support tools to strongly reduce water consumption in a watershed (Allain, Obiang Ndong, et al., 2018).

3.9.2. "Cold" cross-fertilization

A second phase of cross-fertilization between evaluation and assessment occurred during and after stage 7 of the method (analysis and aggregation of the matrix). First, the integrated assessment offered some interpretive keys to better appraise the significance of the discourses behind the evaluation matrix. For example, although the decision-support tool alternative was quite consensual among stakeholders, it appeared that some of them valued water efficiency in itself (more crop per drop, or more euro per drop), while others valued the decrease in water consumption that was expected to result from the gain in water efficiency (a cause-effect relationship that was "confirmed" by simulation results). The lesson learned from the integrated assessment stage led us to highlight the risk of a rebound effect that could undermine this apparent consensus.

Considering the other side of cross-fertilization, the multi-actor evaluation also enriched the IAM. One important lesson was, for example, that stakeholders showed interest in alternatives despite the results of the integrated assessment. The simulation indicated that the “reduced irrigated area” alternative did not improve the Aveyron low-flow hydrology (Allain, Obiang Ndong, et al., 2018) but it was still valued by some stakeholders. A debate emerged about permanent grasslands in upstream areas: some considered that they made no economic sense in a context of declining livestock production, while others saw them as an agro-ecological infrastructure fulfilling multiple “silent” functions for water quantity in small tributaries and for water quality. This debate emphasized important blind spots of the integrated assessment and modelling: flows of small tributaries (computable but not with enough accuracy), water quality indicators (absent and inaccessible with MAELIA for now), and economic indicators for non-field crops (requiring either new model developments to simulate their yield or new indicators to reflect domestic use).

More generally, the comparison between the multi-actor multi-criteria matrix and the integrated assessment conclusions helped to unravel framing effects inherited from the model. Within our framework, we consider it particularly relevant to clarify which alternatives and which issues (rather than which processes) are best supported by the model, in order to clarify the relationship between modelling choices and assessment results. In our case study, the decision-support tool alternative was the one best supported by the model: MAELIA made it possible to simulate most of its expected positive impacts while neglecting rebound effects, which emphasized the potential reduction of water withdrawals with decision-support tools. By contrast, with the alternative of reduced irrigated areas, most of the arguments supporting this alternative (e.g. benefits for small tributaries) could not be translated into simulated indicators, while the arguments against it could be computed (decrease of the field-crop agricultural production, incapacity to diminish the number of water crises). For the other two alternatives, framing effects were less one-directional. This analysis of framing effects at the scale of alternatives was done reflexively once the final discussion was over. Yet it would certainly have been a plus for the collective deliberation if it had taken place earlier.

4. Discussion

Our application of the S²CE framework to water management emphasized its potential to articulate systemic analysis with deliberation and to offer a compressed representation of the problem that hides neither diversity nor heterogeneity. We can even argue that both approaches were not only combined but used in synergy: the integrated assessment based on MAELIA simulations (analytical-systemic) provided us with new understandings that were often incorporated into stakeholders’ argumentation. In turn, the multi-actor multi-criteria (deliberative) evaluations gave us new interpretive keys for analysing the results of our simulations and discussing modelling choices.

Two major lessons can be derived from our experience: simulation and modelling, when used as heuristics for evaluating alternatives, enrich collective deliberation (4.2), and using multi-faceted artefacts offers a way to “weakly integrate” values and knowledge (4.3). Before detailing them, we turn back to the limitations of the method that its implementation revealed.

4.1. Critical assessment of the method

As any methodological option, the S²CE implies trade-offs. A first one concerns the status of modelling. When they evaluate a scenario, stakeholders strive to insert their specific arguments into a common-good rationale. Although evaluation workshops were asynchronous, the way they were framed forced stakeholders to develop a priori justifications for their judgments to anticipate opposing arguments. We observed in many groups (not all) that scientific and expert-based knowledge (embodied in simulation-based indicators) were considered a benchmark for reaching an agreement. Also, our methodology for assessing scenarios led the stakeholders to question the model (Is it reliable enough? Which processes are included and how much simplified are they? Could have we refined calculations? What were the input data? etc.). but not the legitimacy of modelling itself.

Therefore, extending the collective deliberation with IAM seems to have provided more solid ground for narratives falling into “efficiency - oriented” justification type to the detriment of other types of justifications (in terms of “honesty”, “equity”, “aesthetic”, or “domesticity”, for example). This bias is not specific to the use of expert-based modelling and cans also be found in cases of participatory modelling (Barnaud, Le Page, Dumrongrojwatthana, & Trébuil, 2013).

Another important trade-off is about aggregation. In the case study, we decided to perform an aggregation across indicators. This aggregation responded to a voting principle: the judgment that acquired the highest weight was the one displayed in the final matrix, at the criterion level. We left aside other aggregation options such as weighted means or rankings. A weighted mean requires to use quantitative value judgments that express a compensation (Choo, Schoner, & Wedley, 1999). However, in the case study, value judgments were relative to a reference situation and could not fit a discrete cardinal scale. Two judgments “satisfactory improvement” were not equivalent as the extent of the improvement or of the satisfaction was not assessed. The judgment “satisfactory improvement” did not compensate for “displeasing deterioration” and their combination could not equal to “insignificant change”. Non-compensatory aggregation, such as in outranking procedures, would have exhibited other problems of interpretation. Introducing a ranking for scenarios would have made us lose trace of acceptability / unacceptability (e.g. the preferred scenario against one criterion can still be not satisfactory). However, we expected this duality to be more promising to stimulate a debate. The cons were that aggregated results did not provide information about the internal consistency or discrepancy among value judgments attributed by the participants at the indicator level. Interpretation of the results therefore required navigating between aggregated and disaggregated displays of the matrix, and made comparisons less easy.

To sum up, the use of modelling supplies an efficiency-oriented justification regime; and the use of qualitative value judgments expressing acceptability restricts the possibilities for fine-tuned comparisons. In other words, combining deliberative and analytical approaches means that each approach constrains the other. Besides those constraints and biases, the combination remained a fertile exercise, based on which we can derive useful methodological lessons.

4.2. Cautionary and heuristic use of simulation outputs for evaluation

Although the use of models bears unavoidable limitations (e.g. the difficult inclusion of all types of value systems), we consider it a pity not to benefit from the capacities of “hard models” to better learn about the complexities of social-ecological systems, and especially about counter-intuitive effects of changes: cascading effects, emerging effects, feedback loops etc.. The promise of the S^2CE framework is to put such models at the service of collective deliberation and social learning.

In the S^2CE, the model has the status of a tool for generating new knowledge, hopefully of interest to stakeholders, but not of a “hard fact provider”.. MAELIA, for instance is a powerful tool to learn about low-flow dynamics, the agriculture-hydrology interplay, and cross-scale effects from fields to landscape. The other side of the coin is that this high systemic relevance, which relies on academic and technical knowledge, goes with low and/or partial social relevance. This discrepancy was evidenced by the gaps between the IAM results and those of the multi-actor multi-criteria evaluation.. A link is therefore required between the virtual and socially non-relevant world of modelling and the political arena in which problems are discussed. This challenge was partially met with the S^2CE framework.

In the deliberative stage of the framework, we prompted participating stakeholders on using simulations not as quantitative proxies to objectify a decision but as possible (and non-obligatory) bricks to elaborate value judgments. This heuristic status for simulation and modelling fostered intra-group deliberation (about the relevance of indicators, their meaning, their limitations etc.) and, in some cases, contributed to reshaping group preferences. Some participants took pleasure in “moving out of the frame” (i.e. contesting model hypotheses, using new indicators), others appreciated the workshop format as a way to talk with researchers on an equal footing, while others were interested in discovering the simulation results and trying to understand them. In parallel, stakeholders, with their judgments, questions and comments, provided researchers with new material for analysis and new perspectives. A two-way dialogue between researchers and stakeholders emerged, hence lowering the barrier between “hard facts“ and “soft values“ (Funtowicz & Ravetz, 1993).

As a recommendation to go a step further, we offer to clarify framing biases in relation to specific discourses. We aim here not to repeat one more time that models are simplified and partial representations, but to defend the fact that this partiality should be contextualized. Indeed, the partiality of models is generally expressed in absolute terms and at the scale of processes (e.g. the SWAT model does not allow robust estimates of water flows for small tributaries to be produced), which is relevant for a laboratory exercise, but not for collective deliberation. We emphasize that biases and simplifications should be related to the broader discourses of stakeholders, and explained in terms of which arguments the model best supports and which ones are left aside. This claim is totally congruent with the idea that numerical indicators, including simulation outputs, should be embedded within narratives (Saltelli & Giampietro, 2017).

4.3. “Weak integration” and multi-faceted artefacts

Finding a common language has already been pointed out as a necessity in transdisciplinary research and multi-actor problems (Brandt et al., 2013; Ramos et al., 2015). In the case study, some crucial elements of this common language were:

- the criteria grid, constructed through a bottom-up approach, which allowed different stakes to be represented, even those not benefitting from powerful spokespersons or not inscribed in regulatory frameworks and management norms;
- the term “indicator”, which gave authority to the arguments of every group, even the arguments that were “out of the box” of the model or considered at the margins of quantitative water management issues;
- the use of qualitative value judgments (expressed by colours), which certainly frustrated some stakeholders willing to be more precise, but contributed to put softer judgments on an equal footing with technical ones. This choice partly counterbalanced the bias towards efficiency-oriented justifications that modelling produced.

These method choices resulted from our endeavour to deconstruct the ideas that one integrative indicator (the Minimum Flow Requirement) could sum up all water issues and that using technical arguments makes the debate more peaceful and rational. Even out of the conflict-prone context of our case study, we believe that some artefacts conceived as integrators – e.g. decision matrixes and indicators – could support deliberation if they are turned into “weak integrators”, embedding many reading options.

First example: the Kerbabel matrix. . Value judgments are set for all stakeholders but indicators do not need to be shared. Moreover, the resulting multi-actor multi-criteria matrix is common to each reader but, owing to visualization options, there is much leeway for interpretation. The reader can navigate between different levels of aggregation (from synthetic judgments to each indicator and value judgments) and from one side of the cube to another, hence looking either at differences among stakeholders, criteria or alternatives.

Second example: the treatment and processing of simulation outputs. Producing a wide range of indicators, even with apparent redundancy, allows users to look either for general differences among alternatives or for more specific aspects (performance during the “bad” years, temporal distribution, impacts on some specific areas, etc.). Moreover, the use of disaggregated forms for indicators (e.g. maps and boxplots rather than means), which we tried to favour in our indicator booklet, caused users to form different value judgments from the same piece of information (Allain, Plumecocq, et al., 2018). Such recommendations for indicators differ from those generally given for integrated assessments, e.g. scientific relevance or practicability (mostly associated with quantification), which aim at improving the assessment's efficacy for decision-making but with the side-effect of impoverishing sustainability debates.

In short, the S^{2}CE framework allows handling computer models for collective deliberation without the need to create a new model from scratch (by contrast with participatory modelling), which is, first, a fastidious task and, second, marginalizes already-existing knowledge on complex systems' behaviour. The originality of the framework is to adapt not the model but its status and use to deliberation. The suggested adaptations (non-obligatory use of model outputs, diversity of indicators and indicator forms, expression of value judgments in qualitative and acceptability terms, navigation between different levels of aggregation and reading options) grant non-experts with the capacity to learn from integrated models that we would otherwise discard for their low legibility and social relevance.

Conclusion

Although systemic-analytical assessments and deliberative evaluations are not competing approaches to sustainability issues, there are some discreet conceptual differences that end up producing a gap between the two approaches. As usual with interdisciplinary issues, helping hands are lent but do not always reach one another. A reconciliation is therefore needed, as each of those approaches bear complementary strengths: the capacity to envision the counter-intuitive effects of changes affecting complex systems (systemic-analytical assessments), and the capacity to envision the social dilemmas and trade-offs inherent to multi-actor contexts (deliberative evaluation).

The methodological framework called S^{2}CE presented in this article creates a conceptually grounded bridge between these approaches. The final stage of the framework consists in cross-fertilizing the knowledge gained through integrated assessment and deliberative evaluation. Both activities are not independent but, in addition to their combination "on the spot", we wanted to promote a reflexive moment, which actually gives sense to the whole methodology. It is through this cross-fertilization that some assumptions behind apparent social consensus come to light, as well as model framing biases. These precious insights represent the added value of the framework, by comparison with a simple adjunction of assessment tools.

We built the S^{2}CE framework not as a new recipe for addressing sustainability issues but rather as a call for using the capacity of complex computer models outside the walls of laboratories, and for having those models feed, but not drive, collective deliberation in contexts of irreducible uncertainties and value conflicts. More generally, we encourage practices that build on already-existing complex models to deal with urgent sustainability issues, adapt the use of these models (rather than their content) to socially diverse arenas, and integrate them with other methods, oriented towards the articulation of values rather than the production of knowledge, for the elaboration of social choices.

References

Allain, S. (2018). Vers une gestion structurelle de l'eau dans un territoire agricole en tension Une démarche d'évaluation multicritère multi-acteur utilisant des simulations informatiques. Retrieved from http://www.theses.fr/2018INPT0061

Allain, S., Obiang Ndong, G., Lardy, R., & Leenhardt, D. (2018). Integrated assessment of four strategies for solving water imbalance in an agricultural landscape. Agronomy for Sustainable Development, 38(6), 60. doi: 10.1007/s13593-018-0529-z

Allain, S., Plumecocq, G., & Leenhardt, D. (2017). How Do Multi-criteria Assessments Address Landscape-level Problems? A Review of Studies and Practices. Ecological Economics, 136, 282–295. doi: 10.1016/j.ecolecon.2017.02.011

Allain, S., Plumecocq, G., & Leenhardt, D. (2018). Spatial aggregation of indicators in sustainability assessments: Descriptive and normative claims. Land Use Policy, 76, 577–588. doi: 10.1016/j.landusepol.2018.02.038

Allison, A. E. F., Dickson, M. E., Fisher, K. T., & Thrush, S. F. (2018). Dilemmas of modelling and decision-making in environmental research. Environmental Modelling & Software, 99, 147–155. doi: 10.1016/j.envsoft.2017.09.015

Argent, R. M., Sojda, R. S., Giupponi, C., McIntosh, B., Voinov, A. A., & Maier, H. R. (2016). Best practices for conceptual modelling in environmental planning and management. Environmental Modelling & Software, 80, 113–121. doi: 10.1016/j.envsoft.2016.02.023

Barnaud, C., Le Page, C., Dumrongrojwatthana, P., & Trébuil, G. (2013). Spatial representations are not neutral: Lessons from a participatory agent-based modelling process in a land-use conflict. Environmental Modelling & Software, 45, 150–159. doi: 10.1016/j.envsoft.2011.11.016

Bond, A., Morrison-Saunders, A., Gunn, J. A. E., Pope, J., & Retief, F. (2015). Managing uncertainty, ambiguity and ignorance in impact assessment by embedding evolutionary resilience, participatory modelling and adaptive management. Journal of Environmental Management, 151, 97–104. doi: 10.1016/j.jenvman.2014.12.030

Brandt, P., Ernst, A., Gralla, F., Luederitz, C., Lang, D. J., Newig, J., … von Wehrden, H. (2013). A review of transdisciplinary research in sustainability science. Ecological Economics, 92, 1–15. doi: 10.1016/j.ecolecon.2013.04.008

Brugnach, M., Dewulf, A., Henriksen, H. J., & van der Keur, P. (2011). More is not always better: Coping with ambiguity in natural resources management. Journal of Environmental Management, 92(1), 78–84. doi: 10.1016/j.jenvman.2010.08.029

Budds, J. (2009). Contested H2O: Science, policy and politics in water resources management in Chile. Geoforum, 40(3), 418–430. doi: 10.1016/j.geoforum.2008.12.008

Chamaret, A., O'Connor, M., & Douguet, J.-M. (2009). KerDST: The Kerbabel™ on-line deliberation support tool. Centre of Economics and Ethics for Environment and Development. University of Versailles Saint-Quentin-En-Yvelines: France. Retrieved from http://www.publicsphereproject.org/events/diac08/proceedings/06.KerDST.Chamaret_et_al.pdf

Choo, E. U., Schoner, B., & Wedley, W. C. (1999). Interpretation of criteria weights in multicriteria decision making. Computers & Industrial Engineering, 37(3), 527–541.

De Marchi, B., Funtowicz, S. O., Lo Cascio, S., & Munda, G. (2000). Combining participative and institutional approaches with multicriteria evaluation. An empirical study for water issues in Troina, Sicily. Ecological Economics, 34(2), 267–282. doi: 10.1016/S0921-8009(00)00162-2

Debril, T., & Therond, O. (2012). Les difficultés associées à la gestion quantitative de l'eau et à la mise en oeuvre de la réforme des volumes prélevables: Le cas du bassin Adour-Garonne. Agronomie, Environnement et Sociétés, 2(2), 127–138.

Elgert, L. (2013). Hard Facts and Software: The Co-production of Indicators in a Land-use Planning Model. Environmental Values, 22(6), 765–786. doi: 10.3197/096327113X13781997646610

Fernandez, S. (2014). Much Ado About Minimum Flows…Unpacking indicators to reveal water politics. Geoforum, 57, 258–271. doi: 10.1016/j.geoforum.2013.04.017

Frame, B., & Brown, J. (2008). Developing post-normal technologies for sustainability. Ecological Economics, 65(2), 225–241. doi: 10.1016/j.ecolecon.2007.11.010

Frame, B., & O'Connor, M. (2011). Integrating valuation and deliberation: The purposes of sustainability assessment. Environmental Science & Policy, 14(1), 1–10. doi: 10.1016/j.envsci.2010.10.009

Funtowicz, S., & Ravetz, J. (1990). Uncertainty and Quality in Science for Policy. Springer Science & Business Media.

Funtowicz, S., & Ravetz, J. (1993). Science for the post-normal age. Futures, 25(7), 739–755. doi: 10.1016/0016-3287(93)90022-L

Funtowicz, S., & Ravetz, J. (1994). The worth of a songbird: Ecological economics as a post-normal science. Ecological Economics, 10(3), 197–207. doi: 10.1016/0921-8009(94)90108-2

Gaudou, B., Sibertin-Blanc, C., Therond, O., Amblard, F., Auda, Y., Arcangeli, J.-P., … others. (2013). The MAELIA multi-agent platform for integrated analysis of interactions between agricultural land-use and low-water management strategies. International Workshop on Multi-Agent Systems and Agent-Based Simulation, 85–100. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-54783-6_6

Giampietro, M. (2003). Multi-Scale Integrated Analysis of Agroecosystems. CRC Press.

Giampietro, M., Mayumi, K., & Munda, G. (2006). Integrated assessment and energy analysis: Quality assurance in multi-criteria analysis of sustainability. Energy, 31(1), 59–86. doi: 10.1016/j.energy.2005.03.005

Guimarães Pereira, A., & Funtowicz, S. O. (2006). Knowledge representation and mediation for transdisciplinary frameworks: Tools to inform debates, dialogues & deliberations. International Journal of Transdisciplinary Research, 1(1), 34–50.

Haag, D., & Kaupenjohann, M. (2001). Parameters, prediction, post-normal science and the precautionary principle—A roadmap for modelling for decision-making. Ecological Modelling, 144(1), 45–60. doi: 10.1016/S0304-3800(01)00361-1

Hamilton, S. H., ElSawah, S., Guillaume, J. H. A., Jakeman, A. J., & Pierce, S. A. (2015). Integrated assessment and modelling: Overview and synthesis of salient dimensions.

Environmental Modelling & Software, 64, 215–229. doi: 10.1016/j.envsoft.2014.12.005

Jakeman, A. J., & Letcher, R. A. (2003). Integrated assessment and modelling: Features, principles and examples for catchment management. Environmental Modelling & Software, 18(6), 491–501. doi: 10.1016/S1364-8152(03)00024-0

Jasanoff, S. (2003). Technologies of Humility: Citizen Participation in Governing Science. Minerva, 41(3), 223–244. doi: 10.1023/A:1025557512320

Kok, K., Rothman, D. S., & Patel, N. (2006). Multi-scale narratives from an IA perspective: Part I. European and Mediterranean scenario development. Futures, 38(3), 261–284. doi: 10.1016/j.futures.2005.07.001

Leenhardt, D., Therond, O., Cordier, M.-O., Gascuel-Odoux, C., Reynaud, A., Durand, P., … Moreau, P. (2012). A generic framework for scenario exercises using models applied to water-resource management. Environmental Modelling & Software, 37, 125–133. doi: 10.1016/j.envsoft.2012.03.010

March, H., Therond, O., & Leenhardt, D. (2012). Water futures: Reviewing water-scenario analyses through an original interpretative framework. Ecological Economics, 82, 126–137. doi: 10.1016/j.ecolecon.2012.07.006

Martinez-Alier, J., Munda, G., & O'Neill, J. (1998). Weak comparability of values as a foundation for ecological economics. Ecological Economics, 26(3), 277–286. doi: 10.1016/S0921-8009(97)00120-1

Miller, C. A., O'Leary, J., Graffy, E., Stechel, E. B., & Dirks, G. (2015). Narrative futures and the governance of energy transitions. Futures, 70, 65–74. doi: 10.1016/j.futures.2014.12.001

Munda, G. (2004). Social multi-criteria evaluation: Methodological foundations and operational consequences. European Journal of Operational Research, 158(3), 662–677. doi: 10.1016/S0377-2217(03)00369-2

O'Connor, M., Small, B., & Wedderburn, E. M. (2010). Sustainable Agriculture in Aotearoa: Social Learning through Piecewise Deliberation. Retrieved from https://hal.archives-ouvertes.fr/hal-00879693

O'Connor, M., & Spangenberg, J. H. (2008). A methodology for CSR reporting: Assuring a representative diversity of indicators across stakeholders, scales, sites and performance issues. Journal of Cleaner Production, 16(13), 1399–1415.

Paavola, J. (2007). Institutions and environmental governance: A reconceptualization. Ecological Economics, 63(1), 93–103. doi: 10.1016/j.ecolecon.2006.09.026

Pahl-Wostl, C., Schlumpf, C., Büssenschütt, M., Schönborn, A., & Burse, J. (2000). Models at the interface between science and society: Impacts and options. Integrated Assessment, 1(4), 267–280. doi: 10.1023/A:1018927120883

Pahl-Wostl, C., Sendzimir, J., Jeffrey, P., Aerts, J., Berkamp, G., & Cross, K. (2007). Managing Change toward Adaptive Water Management through Social Learning. Ecology and Society, 12(2). doi: 10.5751/ES-02147-120230

Petersen, A. C., Cath, A., Hage, M., Kunseler, E., & van der Sluijs, J. P. (2011). Post-Normal Science in Practice at the Netherlands Environmental Assessment Agency. Science, Technology, & Human Values, 36(3), 362–388. doi: 10.1177/0162243910385797

Ramos, J., Soma, K., Bergh, O., Schulze, T., Gimpel, A., Stelzenmueller, V., … Gault, J. (2015). Multiple interests across European coastal waters: The importance of a common language. Ices Journal of Marine Science, 72(2), 720–731. doi: 10.1093/icesjms/fsu095

Rauschmayer, F., & Wittmer, H. (2006). Evaluating deliberative and analytical methods for the resolution of environmental conflicts. Land Use Policy, 23(1), 108–122. doi: 10.1016/j.landusepol.2004.08.011

Salliou, N., Barnaud, C., Vialatte, A., & Monteil, C. (2017). A participatory Bayesian Belief Network approach to explore ambiguity among stakeholders about socio-ecological systems. Environmental Modelling & Software, 96, 199–209. doi: 10.1016/j.envsoft.2017.06.050

Saltelli, A., & Giampietro, M. (2017). What is wrong with evidence based policy, and how can it be improved? Futures, 91, 62–71. doi: 10.1016/j.futures.2016.11.012

Sardar, Z. (2010). The Namesake: Futures; futures studies; futurology; futuristic; foresight—What's in a name? Futures, 42(3), 177–184. doi: 10.1016/j.futures.2009.11.001

Spash, C. L. (2009). The New Environmental Pragmatists, Pluralism and Sustainability. Environmental Values, 18(3), 253–256. Retrieved from JSTOR.

Stagl, S. (2006). Multicriteria evaluation and public participation: The case of UK energy policy. Land Use Policy, 23(1), 53–62. doi: 10.1016/j.landusepol.2004.08.007

Temper, L., Delbene, D., Martinez-Alier, J., & Rodriguez-Labajos, B. (2015). Mapping the frontiers and front-lines of Environmental Justice: The EJOLT Atlas. Journal of Political Ecology, 22, 255–278.

Tress, G., Tress, B., & Fry, G. (2007). Analysis of the barriers to integration in landscape research projects. Land Use Policy, 24(2), 374–385. doi: 10.1016/j.landusepol.2006.05.001

Tschakert, P., Dietrich, K., Tamminga, K., Prins, E., Shaffer, J., Liwenga, E., & Asiedu, A. (2014). Learning and Envisioning under Climatic Uncertainty: An African Experience. Environment and Planning A: Economy and Space, 46(5), 1049–1068. doi: 10.1068/a46257

Turnpenny, J., Jones, M., & Lorenzoni, I. (2011). Where Now for Post-Normal Science?: A Critical Review of its Development, Definitions, and Uses. Science, Technology, & Human Values, 36(3), 287–306. doi: 10.1177/0162243910385789

van Asselt, M., & Rijkens-Klomp, N. (2002). A look in the mirror: Reflection on participation in Integrated Assessment from a methodological perspective. Global Environmental Change, 12(3), 167–184. doi: 10.1016/S0959-3780(02)00012-2 van der Sluijs, J. P. (2002). A way out of the credibility crisis of models used in integrated environmental assessment. Futures, 34(2), 133–146. doi: 10.1016/S0016-3287(01)00051-9

Vatn, A. (2009). An institutional analysis of methods for environmental appraisal. Ecological Economics, 68(8–9), 2207–2215. doi: 10.1016/j.ecolecon.2009.04.005

Vervoort, J. M., Kok, K., van Lammeren, R., & Veldkamp, T. (2010). Stepping into futures: Exploring the potential of interactive media for participatory scenarios on social-ecological systems. Futures, 42(6), 604–616. doi: 10.1016/j.futures.2010.04.031

Voinov, A., Kolagani, N., McCall, M. K., Glynn, P. D., Kragt, M. E., Ostermann, F. O., … Ramu, P. (2016). Modelling with stakeholders – Next generation. Environmental Modelling & Software, 77, 196–220. doi: 10.1016/j.envsoft.2015.11.016

Walker1a, B., Carpenter, S., Anderies1b, J., Abel1b, N., Cumming, G., Janssen, M., … Pritchard, R. (2002). Resilience management in social-ecological systems: A working hypothesis for a participatory approach. Conservation Ecology, 6(1), 14.

Zografos, C. (2015). Value deliberation in ecological economics. Handbook of Ecological Economics. Retrieved from https://www.elgaronline.com/view/edcoll/9781783471409/9781783471409.00008.xml