Eawag
Überlandstrasse 133
P.O. Box 611
8600 Dübendorf
Switzerland
+41 (0)58 765 55 44
www.eawag.ch

Environmental Social Sciences (ESS)
Dr. Judit Lienert
Cluster Leader Decision Analysis
+41 (0)58 765 55 74
judit.lienert@eawag.ch
https://www.eawag.ch/en/aboutus/portrait/organisation/staff/profile/judit-lienert/show/

Dübendorf, 10. December 2021

Response to referee # 1 and editors

Dear Referee, dear Editors

Thank you very much for reviewing our manuscript:

Judit Lienert, Jafet Andersson, Daniel Hofmann, Francisco Silva Pinto, Martijn Kuller, "Can MCDA guide transdisciplinary endeavors? A framework applied to co-developing a flood forecasting system in West Africa". hess-2021-506

This manuscript was written for the HESS Special Issue **"Contributions of transdisciplinary approaches to hydrology and water resources management"**

We are grateful for the work that has gone into reviewing our paper. We do know that this takes a lot of time, which receives no direct reward. Your suggestions are very constructive and most welcome. We are very willing to improve the manuscript based on your inputs, wherever possible.

We have addressed your comments one-by-one below. *The referees' comments are given in Italics*, our response is given in normal font.

We have a general concern regarding coherence between the two referees. Referee # 1 asks for extending several parts (comments # 3 – 9). In contrast, referee # 2 has "major concerns regarding the text length" and proposes to "eliminate the unnecessary parts of the text so that the necessary ones can 'speak'" (e.g., comments # 3, 8, 27). At the same time, referee # 2 acknowledges that this paper contains a lot of information that could actually be presented in two separate papers (comment # 4).

We agree with referee # 2 that clarification and restructuring could be helpful to increase the papers' understandability and are willing to follow the concrete suggestions, e.g., to rewrite the Abstract and restructure the Introduction (e.g., comments # 3, 4, 5, 6, 8, 22, 26, 28).

Specifically, referee # 2 asks us to follow a traditional set-up for the Introduction (comments # 5, 6, 22, 26). We had not done this, because we are combining literature and research gaps from various fields and because we are following two main aims (as pointed out by referee # 2, comment # 4). However, we made a suggestion for restructuring, and kindly ask the **editors for advice on this; please see referee # 2, comment # 5**.

We emphasize that we do not wish to split the paper into two, because this was what we already had in an earlier version (focusing on the development of the flood forecasting system using MCDA). We think that the more holistic approach encompassing the

transdisciplinary framework and MCDA is beneficial and can raise broader interest, especially to readers that are not specialists (on MCDA and transdisciplinary projects).

Moreover, we propose to not add additional text, e.g., concerning an overview of different MCDA methods (see referee # 1, comment # 3; and referee # 2, comment # 9).

We are still uncertain how we can considerably shorten the paper. Reasons are:

a) The review of literature from transdisciplinary research and sustainability science, is needed as background information and for coming up with a framework.

b) Presenting MCDA is needed for readers of HESS that are likely unfamiliar with an MCDA process. We would not do this in such detail in a Decision Analysis journal. This includes a short review of MCDA in flood management, motivating a typical MCDA process, MCDA method explanations in the Methods section, and a brief overview of main MCDA results in the Results section. We kept this as short as sensibly possible, much additional information is given in the Supplementary Information. Several comments of referee # 1 indicate that more-detailed information of the MCDA method would be appreciated (comments # 3 – 7). We suggest to refrain from this for reasons of length.

c) The discussion is long, because we combine all these aspects. Especially section 4.3 "Suitability of the MCDA process for guiding large transdisciplinary projects (RQC)" is long. We can shorten it, of course, but this would probably make many insights difficult to understand, and we think that the explanations are needed to "add meat" to Table 4; the table was much appreciated by reviewer # 2 (comment # 34).

We kindly **ask the editors for a decision concerning length**, based on our explanations above. Is substantially reducing the length required? If yes, which sections would they want us to substantially shorten or delete?

Moreover, we kindly ask for advice, whether the **editors want us to follow a traditional approach for the Introduction**, which means taking apart the topics that belong to the transdisciplinary framework and those that belong to the MCDA (i.e., first we would present literature from TD research and MCDA, then the research questions, then methods from TD research and MCDA).

We much appreciate feedback. With best regards,

Judit Lienert

also on behalf of my co-authors, Jafet Andersson, Daniel Hofmann, Francisco Silva Pinto, and Martijn Kuller

Anonymous Referee #1

Citation: https://doi.org/10.5194/hess-2021-506-RC1

1) *The draft manuscript provided an interesting approach to designing a transnational/regional flood forecasting system using MCDA. This is a rather new approach to integrating user requirements on systems design considering the large scale and engagement of many transnational stakeholders. Whilst, MCDA has been used in environmental topics such as to assess vulnerability, risk and decision making in finding the best solutions.*

   The approach built upon the concept of inter and transdisciplinary research, integrating stakeholders' expert opinions with scientifically sound facts. The development of the system started with stakeholder analysis to identify the right stakeholders to include in the process. The requirement for the MCDA was then defined in a participatory manner, the objectives/criteria, attributes/indicator and weights.

   In general, the important concepts has been sufficiently described in the manuscript.

Response: Thank you for the overall positive assessment of our work.

2) *There is however a concept that I think was misleading. On page 9 line 233 –" As first step of the MCDA process, we undertook a stakeholder analysis (e.g., Grimble and Wellard, 1997;Lienert et al., 2013;Reed et al., 2009), which is often neglected in MCDA projects." , Page 25, line 577, "..MCDA…it can be very suitable for identifying stakeholders …." and similarly in page 28, line 682-683. In my perspective stakeholder analysis is important when you do any participatory research works and MCDA is a technique to integrate different criteria to select the best possible options. Hence it is not always that MCDA equates to stakeholder identification.*

Response: Thank you; our formulation seems misleading. We are happy to clarify this in the revisions: We fully agree that stakeholder analysis is useful for any type of participatory re-search. By no means did we mean to imply that the method is unique to MCDA. However, as detailed in various parts of the paper, we understand MCDA as an entire (transdisciplinary) process, which includes stakeholders at various stages of this process. We do not regard MCDA solely as a technique to integrate different criteria.

For instance, we wrote on **p. 2, line 57**: "We organized stakeholder participation adopting a comprehensive Multi-Criteria Decision Analysis (MCDA) process (Belton & Stewart, 2002; Eisenführ et al., 2010; Keeney, 1982) that includes problem structuring methods (PSMs; Rosenhead & Mingers, 2001)."

One such problem structuring method is stakeholder analysis. We wish to emphasize that the paper is based on this understanding of MCDA as a transdisciplinary process.

Please see the aims, **p. 3, line 91**: "Taking a more theoretical stance, we aimed to critically analyze FANFAR from the perspective of knowledge co-creation, sustainability science, and transdisciplinary research. We focused on MCDA as a process, rather than a technical method (sect. 2.2.; Figure 1). We set up a framework drawing from literature to uncover strengths and weaknesses of MCDA to guide the transdisciplinary process (RQC)."

3) *To improve the manuscript, I would suggest providing a brief review of different MCDA techniques such as SAW, TOPSIS, etc. and to elaborate on why the choice of the compensatory method used in FANFAR.*

Response: We understand that it can be interesting to provide an overview of different MCDA methods. However, referee # 2 clearly asks us to shorten the paper, so adding additional text seems counter-productive. Referee # 2 addressed your suggestion (comment # 9): *"I disagree with referee 1 that asks for a review of standard MCDA methods (there are several of these out there). On the other hand, I agree with referee 1 that the choice for the compensatory method should be clarified."*

We therefore decided to follow referee # 2 with our answer: "Thank you for supporting a decision that we had made in an earlier version of this manuscript. For length reasons, and because we do not find it necessary, we will not introduce the many different possible MCDA methods. It would entail describing advantages and disadvantages of e.g., AHP, different outranking methods (PROMETHEE, ELECTREE, etc.), and newer approaches such as TOPSIS, additionally to MAVT/MAUT. We cite a classic textbook and a review paper from hydrology that both provide overviews of methods:"

p. 8, line 208: "MCDA embraces various methodologies to support complex decisions (e.g., Belton & Stewart, 2002; de Brito & Evers, 2016)."

Response regarding "compensatory method": We wish to emphasize that we did NOT choose a compensatory method. Rather, MAVT/MAUT is very flexible regarding the mathematical (aggregation) model. It is only fully compensatory if standard linear additive aggregation is used. We describe our non-additive, i.e., non-compensatory approach in the Methods section 2.2.9:

p. 14, line 383: "While easy to understand, the additive model entails strong assumptions, e.g., that objectives are preferentially independent (Eisenführ et al., 2010). Increasing evidence indicates that many stakeholders do not agree with model implications (Haag et al., 2019; Reichert et al., 2019; Zheng et al., 2016). Additive aggregation implies that good performance on one objective can fully compensate for poor performance on another. In the FANFAR weight elicitation sessions, we asked stakeholders (…)."

We agree with both reviewers that this was unclear. We will add this advantage to the **Introduction section 2.2.2, page 8:** MAVT/MAUT provides large flexibility in the mathematical model choice, including non-compensatory aggregation.

---

4)   *Similarly, also for the weighting method. Please also elaborate more clearly on how the value function curve for the attributes was derived. In a participatory manner and what is the process? On page 13 line 347, here it presented that seven evenly spaced levels were created for the sub-attribute (worst, very bad, bad, neutral, good, very good, and best). This attribute level was transformed from 0-1 values using linear interpolation. Having a linguistic term, why was the Fuzzy set theory (Chen and Wang 1992) not considered in converting it to crisp number values?*

Response: We fully understand your request, but are in a dilemma here. Providing more information on elicitation procedures will inevitably increase the length of the paper, which we wish to avoid (see comments by referee # 2 regarding length). For weighting, we used standard methods (see our response to your comment # 5).

We described how we elicited the shape of marginal value functions from experts in the **main text, section 2.2.8**, which covers nearly half a page (**p. 13, lines 338 – 353**). For reasons of space, we think we should not elaborate further on how the value functions were constructed in the main text. We specifically refer to the **Supplementary Information**. Note, in the case of marginal value functions, we did not ask the workshop participants for their preferences. We often do this in workshops in other case studies, usually using the standard bisection

method for elicitation (Eisenführ et al., 2010). However, in the FANFAR case, we were convinced that expert input is required, and that the workshop participants who were not familiar with many of the attributes could not have provided meaningful answers to our elicitation questions. see:

p. 13, line 344: "In FANFAR, most attributes are relatively technical, requiring expert knowledge. We therefore elicited shapes of value functions from experts (sect. 2.2.7; details, including figures of value functions, see sect. SI-2.4.1)."

Moreover, the attributes were complex; many of them consisted of several sub-attributes. We provide an overview in **Figure 2**; the sub-attributes are to the far right. The predictions for each attribute and the corresponding value functions were constructed carefully together with the different experts, who had a deep understanding of the forecast and alert system and the respective attributes. Unfortunately, it is not possible to convey this information in a few simple sentences in the main text.

In the **Supplementary Information**, we provide **26 pages** regarding the attributes and marginal value functions. For each of the 10 attributes, we show in detail on 1 – 2 pages how it was constructed; including about 2 figures and 2 tables for each attribute. This information starts on **p. 26 with section 2.4. "Predicting performance of each system configuration"**, and ends on **p. 52 with section 2.5 "Marginal value functions"**. We think that this information is useful and provides the raw data background that allows verifying our results and replicating them (if ever anybody would be interested in doing so). However, we do not think that this type of raw data information should enter the main text.

Furthermore, you propose using fuzzy set theory to deal with uncertainty. This is one possibility; we chose to work with **probability theory**. Arguments for using probability theory when working with Multi-Attribute Value Theory (MAVT) or Utility Theory (MAUT) are detailed in our paper Reichert et al. (2015). We think it is out of bounds here to go into further details. Regarding the practical information on how we did it: we provided all necessary details regarding uncertainty of the raw data in the aforementioned sections in the **Supplementary Information**. Each attribute has a sub-header "**Uncertainty of predictions**" where the very interested reader can find this information.

---

5) *Page 13, line 355, elaborate why two methods for weighting was used for different language groups. Why Swing and Simos?*

Response: Both Swing and Simos' revised card procedure are standard, well-validated weight elicitation methods, and we provide the according references. Main reason for using two different methods in the language groups is practical. Weight elicitation was run in two parallel sessions in the workshops: one decision analyst (J. Lienert) speaks French, and is most familiar with Swing (or the Trade-off method), which is why Swing was used in the French speaking groups. The other analyst (F. Silva Pinto) speaks English and is most familiar with Simos' card procedure. Additionally, we were interested if results might be different.

A justified criticism could be that using different weight elicitation procedures might lead to systematic biases between the groups, since there is some experimental evidence that different weight elicitation procedures might lead to different weights. We have published a paper on this: Lienert et al. (2016). However, although this can be relevant, in FANFAR it was most important to cover a broad range of possible preferences and test, whether such different preferences might lead to another best performing alternative. As the results of the sensitivity analyses showed, this was not the case: even when using maximal and minimal values of weights elicited from the workshop participants in cases where they were uncertain, the results concerning best-performing system configurations were robust.

Additionally, there is no evidence for a systematic bias occurring due to using different weight elicitation procedures in the French and English speaking groups. One group deviated most strongly from all other groups, namely the French speaking emergency managers (group G1A_EM_F). This group also deviated very strongly from the two other French-speaking groups (G2A_HY_F and G2B_HY_F). Put differently: the two other French-speaking groups G2A and G2B had weights that were more similar to the English speaking groups. We suggest not putting this information in the main text, since it would make it even longer.

6) Page 15, line 403, how did you come up with 1,000 Monte Carlo Simulation? Please elaborate on the combinations

Response:

1'000 Monte Carlo simulation runs is a sufficiently large number to capture also smaller deviations in the results of this type of MCDA. More accuracy regarding the uncertainty over attribute predictions does not necessarily provide additional information about best-performing alternatives (i.e., system configurations). Moreover, 1'000 runs is the maximum available in the ValueDecisions app. We did run the analyses with lower numbers and did not get strongly different results. Nevertheless, we decided to use the maximum 1'000 runs because these gave the best resolution and clearest image of the differences between alternatives.

Please see **Figure 6**, which indicates that 1'000 runs suffice to decide which system configurations systematically achieve the top ranks. We reported this result on **p. 6, lines 529 – 533**. Because the weights enter the MCDA model on equal footing, these tend to have a stronger influence on the results. In our case, sensitivity analyses indicated that results were especially sensitive to the aggregation model (**see Table 3**), but again this was insufficient to provoke many rank reversals regarding the three best-performing alternatives.

7) Page 21, line 514, were you able to capture in context why the differences in weight preference? This may give you additional insight into the stakeholders' perspectives.

Response: Thank you for this important observation. Indeed, we were highly interested in this stakeholder perspective. We were able to discuss with the workshop participants, the hydrologists and emergency managers, why the objectives were particularly important or unimportant to them. We provided some of this information in the main text, e.g.:

Results, p. 20, line 499: "Again, the French speaking emergency managers (G1A) were exceptional in assigning much lower weights to objectives they considered unimportant (objectives 23, 31, 41, and 43). They argued that the goal in emergencies is to save lives, and FANFAR system development should focus on achieving fast access to flood alerts (*22_timely_info*; 0.21) and on personnel that can deal with this information (*42_human_resour*; 0.25)."

Discussion, p. 26, line 613: "All groups regarded several languages as unimportant in weight elicitation, despite emphasizing in plenary discussions that language diversity is crucial. When asked to make trade-offs between accuracy and language, they were willing to give up the latter. They were also willing to trade-off higher operation and maintenance costs (except group G1A) and development time in return for receiving a functioning, precise system."

For reasons of space, it was not possible to provide all details of the discussions in the main text. We provided it in the **Supplementary Information: p. 8, Table SI-3**. For each group,

we summarized where they agreed (and why, if they spoke about this), and where there were reasons for disagreement. This informed the sensitivity analyses; see main text:

p. 14, line 362: "We took the mean as main weight and considered strong deviations (difference in weights > 0.2 compared to mean) in later sensitivity analyses (sect. 2.2.10). For Simos' card procedure, two additional weight sets were used, resulting from eliciting a range for one variable. The moderator recorded important comments to inform the sensitivity analyses (Table SI-3)."

8) *Page 33, line 791, many were satisfied with its performance during the rainy days of the year 2020. Were you able to gain some information on some numbers of true and false predictions? This may be helpful to correlate with the satisfaction of experts.*

Response: Thank you for this important remark. We are currently carrying out a systematic daily reforecasting experiment covering 1991 – 2020 for five different model configurations. The experiment is still ongoing, but when results are available, we will publish them and try to link them to the expert satisfaction as suggested. We will shortly add this information to the Discussion section in the revisions.

9) *Page 33, line 805, any discussions on how the FANFAR Flood forecasting system be maintained in the far future?*

Response: This is indeed a very important question for the practical future of the FANFAR forecast and alert system. We can include some short information in the Discussion section around the main strategy for future sustainability (including capacity development, transfer of responsibilities, ICT, and financing), institutional context, and current status (several proposals and dialogues with potential financiers have been held, and some partial support have been secured). However, we wish to point out that this topic is out of scope of this paper, and will make even longer, which is very much opposed to the comments by reviewer # 2.

10) *I also find that the drafted manuscript needs revisions on the flow of thought in writing. Please see across the manuscript. Numerous contexts are somehow incoherent. Such examples were:*

page 2, line 31, "Worldwide, good operational flood forecast systems, giving accurate, timely, precise, and understandable forecast information and alerts, provide effective and affordable help to anticipate and minimize flood impacts (Perera et al., 2019)".

page 29 line 707 " When creating FANFAR system configurations, it became evident that e.g., frequent power cuts and slow internet in West Africa need consideration. Multi-Attribute Value Theory (Eisenführ et al., 2010) allows later including system configurations (Reichert et al., 2015). ".

Page 27, line 660, I do not see the relevance on mentioning the MCDA in building collaborative research "…In the FANFAR project, building the collaborative research (or project) team with consortium partners from Europe and West Africa was achieved (step 1a, Table 1), but not by MCDA…" The statement somehow does not fit.

Response: Thank you for helping us to improve the papers' coherence. We will reformulate these sentences when revising the paper. We will check the entire text again.

Other comments:

11) *For some parts, the literature, section and abbreviation/text referencing are cited almost every after sentence. This becomes inconvenient to read. Such example is page 42, line 688 "To foster joint understanding, commitment, and trust, many of 31 analyzed transdisciplinary projects provided e.g., trainings, or attractive visualizations of recent research (Schneider et al., 2019). Capacity building can be promoted by working in integrated ways of knowledge coproduction discussed above (Caniglia et al., 2021), or with capacity building courses (Wuelser et al., 2021). The FANFAR project offered many training and capacity building opportunities, which cannot be attributed to MCDA."*

Response: We understand that reading the text with many literature citations is bumpy. However, we do find it important to cite the appropriate references in the correct places, especially since they come from different fields that are less familiar to most readers of HESS. The examples above stem from the transdisciplinary literature. We think that this unavoidable to some extent. However, we are willing to carefully check the text again during revisions and remove citations that may be less central.

12) *Page 18, line 468, table 2 text description can be presented as a footnote below the table rather than as part of the caption. Same for table 3 and others*

Response: We understood that HESS asks to avoid footnotes wherever possible, and we tried to follow this requirement.

13) *Page 19 "..g_Attractve Most attractive in West Africa:…" This I do not understand.*

Response: Thank you for pointing this out. We will re-write the sentence for clarification.

14) *Page 24 Figure 7. I see no value of information of having colors of dots.*

Response: The colors of the dots represent unique respondents, enabling a comparison of answers across the three questions for each respondent. We attempted to explain this in the figure caption: "Colored dots represent unique respondents (N = 12)".

References

Belton, V., & Stewart, T. S. (2002). *Multiple Criteria Decision Analysis: An Integrated Approach*. Springer. https://doi.org/10.1007/978-1-4615-1495-4 (Originally published by Kluwer Academic Publishers in 2002)

de Brito, M. M., & Evers, M. (2016). Multi-criteria decision-making for flood risk management: a survey of the current state of the art. *Natural Hazards and Earth System Sciences*, *16*(4), 1019-1033. https://doi.org/10.5194/nhess-16-1019-2016

Eisenführ, F., Weber, M., & Langer, T. (2010). *Rational Decision Making* (Vol. 1st ed.). Springer.

Haag, F., Lienert, J., Schuwirth, N., & Reichert, P. (2019). Identifying non-additive multi-attribute value functions based on uncertain indifference statements. *Omega-International Journal of Management Science*, *85*, 49-67. https://doi.org/10.1016/j.omega.2018.05.011

Keeney, R. L. (1982). Decision-Analysis - an Overview. *Operations Research*, *30*(5), 803-838. https://doi.org/https://doi.org/10.1287/opre.30.5.803

Lienert, J., Duygan, M., & Zheng, J. (2016). Preference stability over time with multiple elicitation methods to support wastewater infrastructure decision-making. *European Journal of Operational Research*, *253*, 746-760. https://doi.org/http://dx.doi.org/10.1016/j.ejor.2016.03.010.

Reichert, P., Langhans, S. D., Lienert, J., & Schuwirth, N. (2015). The conceptual foundation of environmental decision support. *Journal of Environmental Management*, *154*, 316-332. https://doi.org/10.1016/j.jenvman.2015.01.053

Reichert, P., Niederberger, K., Rey, P., Helg, U., & Haertel-Borer, S. (2019). The need for unconventional value aggregation techniques: experiences from eliciting stakeholder preferences in environmental management. *EURO Journal on Decision Processes*, *7*(3), 197-219. https://doi.org/10.1007/s40070-019-00101-9

Rosenhead, J., & Mingers, J. (2001). *Rational Analysis for a Problematic World Revisited: Problem Structuring Methods for Complexity, Uncertainty and Conflict* (J. Rosenhead & J. Mingers, Eds. Vol. 2nd Edition). John Wiley & Sons.

Zheng, J., Egger, C., & Lienert, J. (2016). A scenario-based MCDA framework for wastewater infrastructure planning under uncertainty. *Journal of Environmental Management*, *183, Part 3*, 895-908. https://doi.org/http://dx.doi.org/10.1016/j.jenvman.2016.09.027