# Can MCDA guide transdisciplinary endeavors? A framework applied to co-developing a flood forecasting system in West Africa

Judit Lienert[1], Jafet C.M. Andersson[2], Daniel Hofmann[1], Francisco Silva Pinto[1], Martijn Kuller[1]

[1]Eawag: Swiss Federal Institute of Aquatic Science & Technology, Environmental Social Sciences (ESS) Department, Ueberlandstrasse 133, 8600 Duebendorf, Switzerland
[2]Swedish Meteorological and Hydrological Institute (SMHI), Hydrology Research, 601 76 Norrköping, Sweden

*Correspondence to*: Judit Lienert (judit.lienert@eawag.ch)

**Abstract.** Climate change is projected to increase flood risks in West Africa. The EU Horizon 2020 project FANFAR co-designed a pre-operational flood ~~forecasting and alert~~early warning system (FEWS) for West Africa in ~~four~~ workshops with 50–60 stakeholders from 17 countries, adopting ~~a~~ Multi-Criteria Decision Analysis (MCDA)~~ process~~. We aimed at: (i) designing a FEWS with West African stakeholders using MCDA; and (ii) evaluating participatory MCDA as a transdisciplinary process. To achieve aim (i), we used MCDA problem structuring and preference elicitation methods in workshops: stakeholder analysis, creating 10 objectives to be achieved by the FANFAR FEWS and 11 possible configurations. Experts predicted FEWS configuration performance, which we integrated with stakeholder preferences in MCDA, and tested in sensitivity analyses. ~~Firstly, we aimed to find a robust configuration of the FANFAR system. We document empirical evidence of MCDA, including stakeholder analysis, jointly creating 10 objectives, and 11 FANFAR system configurations.~~ Three FEWS showed good performance, despite uncertainty, and were robust across different preferences. ~~Stakeholders~~ For stakeholders ~~found~~ it was most important that the ~~system~~ FEWS produces accurate, clear, timely, and accessible flood risk information~~, well before floods~~. ~~Monte Carlo simulation and sensitivity analyses helped identifying three configurations that were robust despite uncertainty of expert predictions and different stakeholder preferences, elicited in group sessions.~~ To achieve aim (ii), ~~Secondly, we investigated if problem structuring helps focus early technical system development. Although partly achieved, full MCDA was necessary to provide convincingly robust configurations. Thirdly,~~ we reviewed sustainability science and transdisciplinary research literature for common characteristics. Our framework emphasizes uncertainty and integrates interdisciplinary knowledge, which are crucial to the earth systems sciences. MCDA can well address both. We critically evaluated MCDA following the framework. The strength of MCDA lies in step 2: co-producing knowledge with stakeholders, providing a consistent methodology, and unambiguous, shared results. Participatory MCDA including problem structuring can contribute to step 1 of co-design, but does not well achieve step 3 of co-disseminating and evaluating results. We encourage colleagues to use the framework for guiding transdisciplinary hydrology research that engages with stakeholders and society.

~~critically analyzed MCDA based on literature from sustainability science and transdisciplinary research. Our proposed framework consists of three steps: co-design (joint problem framing), co-production (doing research), and co-dissemination and evaluation of integrated knowledge. MCDA met many requirements, but not all. In step 1, participatory MCDA with problem~~

structuring provides a consistent methodology, and can identify stakeholders and shared objectives to foster joint understanding. MCDA successfully contributes to step 2 by combining interdisciplinary expert knowledge, integrating conflicting stakeholder preferences, handling uncertainty, and providing unambiguous, shared results. Many elements of step 3 are not met by MCDA. We discuss this framework and using MCDA for transdisciplinary hydrology research that engages with stakeholders and society.

# 1 Introduction

## 1.1 Floods in West Africa

West Africa is vulnerable to the projected impacts of climate change, particularly related toconcerning runoff quantities (Aich et al., 2016;Roudier et al., 2014). While the mechanisms and Climate change projections and mechanisms remain uncertain for West Africa, but there is growing evidence for increased frequency, magnitude, and impact of fluvial floods (Nka et al., 2015). West Africa is already heavily impacted by floods. Preliminary data from the United Nations data estimate that 465 people died from floods in West and Central Africa in 2020. More than 1.7 million people were affected (double the number of 2019), 94'000 people were displaced, and 152'000 houses were destroyed (OCHA, 2020). Worldwide, gGood operational flood early warning system (FEWS)flood forecast systems help minimizing flood impacts (Perera et al., 2019); good means they, giveing accurate, timely, precise, and understandable forecast information, and alerts, provide effective and are affordable help to anticipate and minimize flood impacts (Perera et al., 2019). Several systems FEWS have been set up in for different West African regions, some being very useful. However, none seem to sufficiently meet stakeholder needs regarding: i) timeliness (e.g., annual frequency of PRESASS/PRESAGG forecasts; WMO, 2021); ii) coverage (systems propagating streamflow measurements cover small parts of West Africa and no ungauged basins, e.g., SLAPIS, OPIDIN, FEWS-Oti; Massazza et al., 2020); iii) up-to-date operational production because many models are used only for research (many models are used only for research, e.g., Aich et al., 2016;Schuol et al., 2008), and sometimes systems failwithout failures due to e.g., interrupted data flows or server failures (e.g., interrupted production and access to SATH-NBA during the major 2020 floods; NBA, 2020); iv) accuracy (e.g., global modelling systems such as GloFAS; Passerotti et al., 2020); and v) openness and ownership (e.g., proprietary closed--source consultancy systems, which may limit the capacity and independence of West African scientists and practitionersstakeholders, and hence the systems' FEWS' long-term sustainability). An overview of gaps, needs, and recommendations is provided by WMO (2020). Moreover, stakeholders assigned the lowest score to the overall effectiveness of FEWS in all but one West African country (Lumbroso et al., 2016).

## 1.2 The Developing a FEWS with stakeholders in the FANFAR project

The EU Horizon 2020 project FANFAR (2018–2021) addressed these gaps , running from 2018 to 2021by co-developing a pre-operational FEWS for West Africa addresses these gaps (FANFAR, 2021). It included African and European FANFAR operates in a transnational and transcontinental context, with cconsortium partners, across Europe (Italy, Spain, Sweden, and

~~Switzerland) and Africa (organizations AGRHYMET and NIHSA)~~and~~. It includes~~ stakeholders across 17 countries of West and Central Africa. FANFAR ~~establishes and reinforces existing~~relied on cooperation between ~~European and West African~~ hydrologists, information and communication technology (ICT) experts, decision analysts, and ~~end~~ users ~~such as West African hydrologists and emergency managers~~ (Andersson et al., 2020a)~~ to co-develop a pre-operational flood forecasting and alert system at West Africa scale (“FANFAR system”)~~. ~~It~~ The FANFAR FEWS is currently based on three open-source hydrological HYPE models~~, Niger-HYPE, West Africa HYPE, and World-Wide HYPE~~ (Andersson et al., 2017;Arheimer et al., 2020)~~,~~ ~~employed~~ in a cloud~~-based~~ ICT environment. It includes ~~The d~~daily ~~forecasting chain includes~~ meteorological ~~reanalysis and~~ forecasting, data assimilation~~ of gauge observations and satellite altimetry~~, hydrological ~~initialization and~~ forecasting, flood alert derivation, and distribution through Email, SMS, API, and an ~~web-based ~~i~~I~~nteractive v~~V~~isualization p~~P~~ortal (~~IVP,~~ https://fanfar.eu/ivp/). ~~In this paper, we do not address the technical system and refer interested readers to Andersson et al. (2020b). Rather, we emphasize the complex development(Sultan et al., 2020)~~Rather than the technical system (Andersson et al., 2020b), this paper addresses stakeholder engagement in an iterative co-design process, which is needed to address FEWS development (Sultan et al., 2020).~~ of the FANFAR system in an iterative co-design process, the necessity of which has been recently underlined by Sultan et al. (2020).~~

To organize such a transdisciplinary endeavor involving many stakeholders, ~~At its core are three one-week workshops carried out in West Africa from 2018–2020, each with 50–60 participants, and online workshops in 2021. We organized stakeholder participation adopting a~~a comprehensive Multi-Criteria Decision Analysis (MCDA) process can be suitable (Belton and Stewart, 2002;Eisenführ et al., 2010;Keeney, 1982). It should ~~that includes~~include problem structuring methods (Rosenhead and Mingers, 2001). Participatory MCDA can help focus FEWS development such that it best meets stakeholder expectations. ~~Using MCDA, we integrated the stakeholders’ preferences in the expert assessment of how well different FANFAR system configurations meet stakeholder objectives. Ideally, this helps focus system development on configurations that best meet expectations, given the possibility of contradictory interests concerning the importance of objectives, and uncertainty about how well a system configuration can achieve the objectives.~~Indeed, MCDA has been used in flood risk management (reviewed by de Brito and Evers, 2016;Abdullah et al., 2021), but rarely ~~focusing on a~~as participatory ~~stakeholder~~ process. Stakeholders were not even mentioned in a review of 149 papers (Abdullah et al., 2021). ~~The other review~~ de Brito and Evers (2016) concluded that stakeholder participation was fragmented, despite being reported in 51% of 128 papers, i.e., stakeholders ~~only assigned importance~~ weighted~~s to~~ objectives, but were not involved in the entire decision process~~ (de Brito and Evers, 2016)~~. ~~Participatory~~ However, participation ~~processes involving researchers and non-scientific actors are~~is extensively discussed in ~~literature from~~ sustainability sciences (e.g., Caniglia et al., 2021;Norstrom et al., 2020) and transdisciplinary research (e.g., Jahn et al., 2012;Lang et al., 2012;Mauser et al., 2013;Schneider et al., 2019;Wuelser et al., 2021). With this paper we combine the two fields. We draw from ~~this~~ literature to develop an evaluation framework, and ~~emphasizing continuous stakeholder participation to achieve societal transformation. We~~ apply it to the integrative ~~use~~ MCDA ~~as integrative methodological~~ process in FANFAR~~, which in flood risk management seems to have been mainly used as a technical method, without broader~~

~~stakeholder participation (de Brito and Evers, 2016)~~. To the best of our knowledge, we are not aware of systematic assessment of MCDA ~~processes~~ from the angle of transdisciplinary sustainability research.

## 2 Literature review, research questions

### 2.1 S~~S~~ustainability science and transdisciplinary research frameworks

Disaster management increasingly acknowledges that ~~developing early warning systems~~FEWS development should closely involve users to increase ~~the systems'~~its usefulness (adapted to ~~user~~ needs) and effectiveness (e.g., enhance uptake; Basher, 2006;Bierens et al., 2020;UNISDR, 2010). This follows a trend calling for "collaborative action", "collaborative governance", and "co-production of knowledge"~~ in many fields, e.g., healthcare (Donetto et al., 2015)~~. Sustainability science stresses that societal transformation is needed to address ~~today's~~ global environmental challenges. ~~This~~ It requires close engagement of academia with ~~societal actors~~stakeholders ~~to co-produce and integrate knowledge~~ (e.g., Caniglia et al., 2021;Lang et al., 2012;Mauser et al., 2013;Norstrom et al., 2020;Schneider et al., 2019;Wuelser et al., 2021). Transdisciplinary research emphasizes collaboration between scientific disciplines~~ between different scientific disciplines~~, and ~~between researchers and~~ practitioners (Jahn et al., 2012). It ~~seeks to~~ supports societal problem solving with "situated knowledge" for a ~~specific~~ problem in its socio-ecological context (Wuelser et al., 2021). Many ~~collaborative,~~ transdisciplinary projects ~~involving non-scientific stakeholders~~ are being carried out. However, there is ~~still~~ a lack of systematic integration and conceptualization of empiric evidence (e.g., Lang et al., 2012;Caniglia et al., 2021), and ~~, and of understanding the extent and~~ mechanisms of ~~advancing~~ sustainability transformations are not well understood (Schneider et al., 2019;Wuelser et al., 2021).

Frameworks for collaborative governance ~~usually~~ consist of stages such as inputs, research processes, direct outputs, and further outcomes (Schneider et al., 2019). Earlier ~~conceptual frameworks stem from~~ transdisciplinary ~~literature~~frameworks emphasized~~. Main phases in such~~ iterative processes in phases~~are~~: (i) forming a common research object (see Figure 1 in Jahn et al., 2012), problem framing, ~~and~~ team building (see Figure 1 in Lang et al., 2012); (ii) co-creating solution oriented ~~and transferable~~ knowledge through collaborative research; (iii) applying ~~the~~ co-produced knowledge (Lang et al., 2012), and evaluating its contribution to societal and scientific progress (Jahn et al., 2012). Similar ~~three~~ steps were proposed ~~by an international initiative, where the co-creation of knowledge for~~for ~~addressing~~ global sustainability problems~~ was divided into~~, ~~: (i) co-design; (ii) co-production; and (iii) co-dissemination of knowledge,~~ involving academia and stakeholders throughout (see Figure 3 in Mauser et al., 2013):~~.~~ (i) c~~C~~o-design ~~includes~~with jointly framing ~~of~~ the societal sustainability challenges ~~faced by society~~, research definition, and implementation~~ (e.g., funding)~~; (ii)~~.~~ c~~C~~o-production, ~~-~~includ~~ing~~es methodologically consistent scientific integration of interdisciplinary knowledge~~,~~ ~~including~~and ~~ways for~~ dealing with ~~the~~ uncertainty~~ of results~~; and (iii) ~~. The co-produced knowledge should be societally relevant, which is ensured by continuous interaction with stake-~~

~~holders. ~~c~~C~~o-dissemination ~~of results includes~~via scientific publications and ~~and ~~products ~~that are accessible and understandable by~~for different societal groups, transparent discussion of~~.~~ ~~There should be transparent discussion of r~~results, especially among groups of conflicting interests, and consequential action~~ should be taken~~.

~~Two recent articles in "Nature Sustainability" propose~~Recent frameworks shar~~ed~~ing similar elements, but ~~with ~~used a different structu~~re~~ring. Caniglia et al. (2021) suggest~~ed~~ that research actions for sustainability create transformative change in three knowledge dimensions:~~-~~ (i) ~~knowledge ~~informing intentional design; (ii) ~~knowledge ~~enhancing shared agency by involving multiple actors; and (iii) ~~knowledge ~~enabling contextual realization in ~~constantly ~~changing environments. The dimensions~~se~~ are further characterized by prescriptive knowledge (recommendations ~~about ~~for ~~more desirable~~better options), co-produced knowledge (~~through ~~collaborati~~on~~ve ~~engagement ~~with actors, incorporating their ~~diverse ~~perspectives and interests), and situated knowledge (tailored to specific contexts; see Table 2 in Caniglia et al., 2021). ~~In~~From a ~~more ~~practical perspective, four guiding principles ~~-~~for evaluating the quality and success of co-production processes were proposed~~ by 36 sustainability researchers , underpinned with examples ~~(see Figure 1 in Norstrom et al., 2020)~~. These are~~: (i) context-based (situate process in ~~particular ~~context, place, or issue); (ii) pluralistic (recognize multiple ways of knowing and doing); (iii) goal-oriented (articulate clearly defined, shared goals); and (iv) interactive (ongoing learning among actors, active engagement, and frequent interactions).

Recently, 12 transdisciplinary projects ~~studies systematically evaluated transdisciplinary projects ~~were systematically evaluated to identify common characteristics~~ that are re-usable in other cases~~, resulting in s~~ (Wuelser et al., 2021). Analyzing 12 Swiss transdisciplinary research projects, s~~even types of transferable knowledge ~~were proposed ~~(see Table 2 in Wuelser et al., 2021): (i) transdisciplinary principles (e.g., take practitioners on board~~ from the start~~); (ii) transdisciplinary approaches (~~procedural, ~~e.g., joint problem identification, ~~strategies for building ~~alliances with regional partners); (iii) systematic procedures (e.g., specific methodologies); (iv) product formats (communicate and use ~~research ~~results in practice, e.g., capacity building); (v) experiential know-how (personal learnings, skills~~, and experiences~~); (vi) framings (definitions, descriptions~~ of issues~~); and (vii) insights, data, and information (results~~, outputs~~). Similarly, 31 ~~international ~~transdisciplinary projects were systematically analyzed ~~for ~~to identify three ~~ing ~~generic mechanisms of ~~creating ~~societal impacts (see Figure 1 and Table 2 in Schneider et al., 2019)~~. Three distinct mechanisms and seven strategies were identified~~: (i) promote systems, target, and transformation knowledge~~ (by providing 1. access to information; 2. advice and training; 3. public debates)~~; (ii) foster social learning for collective action~~ (4. multi-stakeholder groups; 5. North-South partnerships)~~; and (iii) enhance competences for reflective leadership~~ (6. transformative education; 7. communities of practice)~~.

The last example stems from the earth systems sciences. Its~~,~~ ~~us~~starting point is~~ing~~ a ~~case study for ~~climate assessment case that ~~as starting point for~~provides conceptual insights, rather than firstly drawing from social science theories as most above ~~examples ~~literature (Lemos and Morehouse, 2005). ~~These authors argue for i~~Hereby, interactive models ~~(= frameworks) ~~and iterative processes foster~~to~~ ~~increase~~ innovation and societal impact across three dimensions: (i) interdisciplinarity (effort of scientists from different disciplines to tackle complex problems, working together iteratively, ~~but~~ or separately if ~~necessary~~needed); (ii) interaction with stakeholders (e.g., problem definition, testing and disseminating results, adapting research to user~~s'~~ needs, building trust); and (iii) production of usable knowledge.

*In this paper, we draw on the reviewed literature from sustainability science and transdisciplinary research to propose a framework for guiding and evaluating transdisciplinary projects. We specifically focus on requirements of hydrology research.*

## 2.2 Multi-Criteria Decision Analysis (MCDA) in flood risk research

Methodological and epistemological perspectives in transdisciplinary research may ~~still~~ be debated (Jahn et al., 2012), but in a project, ~~there needs to be~~ consensus on methods ~~used in a project,~~ and integrative ~~the~~ concepts ~~to integrate research~~is needed (Lang et al., 2012). The FANFAR consortium agreed on MCDA as organizing framework to ~~integrate~~ include ~~West African~~ stakeholders ~~in the transdisciplinary process~~ and ~~achieve the aim of produc~~design~~ing~~ the FEWS for West Africa. MCDA is well suited to address this challenge and embraces various methodologies to support complex decisions (e.g., Belton and Stewart, 2002;de Brito and Evers, 2016). We chose Multi-Attribute Value Theory (MAVT; Eisenführ et al., 2010;Keeney, 1982) for reasons well documented in literature: (i) developing a complex FEWS requires many decisions such as identifying hydrological models and data sources to produce forecasts, or appropriate flood hazard thresholds, visualizations, and distribution channels to reach people. MCDA allows addressing such choices. (ii) To adapt the FEWS to stakeholder needs, collaboration with nonacademic partners is required. MCDA allows close stakeholder interaction, offering various methods for each stage of decision making (e.g., Eisenführ et al., 2010;Keeney, 1982;Marttunen and Hamalainen, 2008;Zheng et al., 2016;Marttunen et al., 2017). (iii) MAVT and Value Focused Thinking (Keeney, 1996) base decisions on the objectives that are of fundamental importance to stakeholders. (iv) To evaluate FEWS configurations, MCDA allows integrating different kinds of scientific and technical data from experts (e.g., forecast accuracy, development costs) with stakeholder preferences. Especially in case of conflicting interests, it can be helpful to disentangle stakeholder values from facts (Gregory et al., 2012a;Keeney, 1982). In complex decisions, not all objectives can be fully achieved. MCDA explicitly asks stakeholders which trade-offs they are willing to make. (v) MAVT and Multi-Attribute Utility Theory (MAUT) are mathematically very flexible. Usually linear additive aggregation is applied, but many non-compensatory models are possible, which may better represent stakeholder preferences (Haag et al., 2019a;Reichert et al., 2015;Reichert et al., 2019). (vi) MAVT/MAUT allow including various types of uncertainty, e.g., of expert predictions with probability theory, or stakeholder preferences with sensitivity analyses (Reichert et al., 2015;Haag et al., 2019b;Zheng et al., 2016). (vii) MCDA is done stepwise to reduce complexity and increase transparency.

MCDA is increasingly popular in hydrology and ~~specifically related to~~ flood risk~~s~~ research. ~~A~~ Our brief literature search revealed around 50 ~~potentially relevant~~ articles, but only few ~~actually~~ included stakeholders (Web of Science ~~search~~ 25.08.2021; keywords: “MCDA” AND “hydrolog*” AND/OR “flood*”). This corroborates results of two ~~recent~~ reviews (de

Brito and Evers, 2016;Abdullah et al., 2021). Both confirmed a significant growth in MCDA applications, especially for flood mitigation, while flood preparedness, response, or recovery phases were understudied. Moreover, Most papers lacked uncertainty analysis and stakeholder participation were rarely applied, and require future exploration to increase the quality of decision making and success of implementation (de Brito and Evers, 2016). Our search indicatesWe found that most papers used MCDA was mainly used as a technical method to (mathematically) integrate - over various indicators, e.g., for to support calibratingon of flood forecasting models (Pang et al., 2019). Recent Recent methodologically interesting papers addressed MCDA coupled with e.g., artificial intelligence (Pham et al., 2021), machine learning (Nachappa et al., 2020), or portfolio decision analysis (Convertino et al., 2019). Combining GIS with MCDA is a trend, also in hydrology, as in other fields. Examples include flood risk assessment explicitly including focusing on uncertainty (Tang et al., 2018), and flood risk analyses a method to producinge risk maps for a Swiss region (e.g., Ronco et al., 2015;Samanta et al., 2016), and flood risk analyses and mappings for specific cases (Samanta et al., 2016).

This applied literature seems to lacks stakeholder integrationing stakeholders in the decision analysis process, even in applied cases, a main conclusion of the review by (de Brito and Evers, 2016). These authors also proposeExceptions are aa -MCDA concept for inter- and transdisciplinary approaches to improve urban resilience in flood risk management using MCDA (Evers et al., 2018), and present a participatory case study in Brazil for flood vulnerability assessment (de Brito et al., 2018). In cases with participation, stakeholders mainly only assigned weights (de Brito and Evers, 2016). Our search confirmed lacking stakeholder involvement. The few exceptions shortly presented weights, without discussing the participatory process (e.g., Ronco et al., 2015). Several However, several studies papers stated that MCDA results are highly susceptible to model assumptions, especially the weights assigned to objectives (de Brito and Evers, 2016). For instance, the sensitivity of MCDA results to weight variabilitvariability was assessed with global sensitivity analysis y was determined with global sensitivity analysis (Tang et al., 2018). In some of our screened papers, stakeholders actively participated, but again (de Brito and Evers, 2016)mainly to assign weights. However, most papers we checked seem technical, and stakeholder involvement was usually described in few sentences, presenting the weights in a table, without further elaboration of the process (e.g., Ronco et al., 2015). Concluding our short review, there seems to be a lack of research in hydrology and flood risk research lack literature that understands MCDA as a process, and that integrates stakeholder engagement processs throughout this process. To increase decision making quality and implementation success, MCDA applications require uncertainty analysis and stakeholder participation (de Brito and Evers, 2016).

*Here, we document participatory MCDA that involves stakeholders throughout the process, including uncertainty. Taking the lens of sustainability science and transdisciplinary research, we apply the proposed framework to critically evaluate MCDA. MCDA embraces various methodologies to support complex decisions (e.g., Belton and Stewart, 2002;de Brito and Evers, 2016). We chose MCDA, and specifically Multi-Attribute Value Theory (MAVT; Eisenführ et al., 2010;Keeney, 1982) for several reasons: (i) to develop a complex forecast system, many decisions had to be made. We needed to clarify, which hydrological models and data should be used to produce flood forecasts (e.g., type of meteorological forecast data, streamflow*

~~*gauge observations), which flood hazard thresholds were appropriate, how to visualize model forecasts, or which distribution channels effectively reach people. MCDA is optimally suited to address such questions. (ii) Because the FEWS should be adapted to stakeholder needs, its development relied on close collaboration with nonacademic partners. MCDA is a proven methodology that allows close interaction with stakeholders at various stages of decision making processes, offering a choice of methods for each stage. (iii) We used MAVT (Keeney, 1996)because it bases decisions on the objectives that are of fundamental importance to stakeholders. (iv) To evaluate options (i.e., system configurations), MCDA allows comparing and integrating different kinds of scientific and technical data from expert predictions (e.g., accuracy of forecasts, estimated costs of system development) with stakeholders' preferences. Disentangling facts from values can be very helpful, especially if stakeholders have conflicting interests (Gregory et al., 2012a;Keeney, 1982). In most complex decisions, not all objectives can be fully achieved. MCDA explicitly asks stakeholders for the trade-offs they are willing to make. (v) MCDA allows including various types of uncertainty, e.g., of expert predictions with probability theory, or uncertain stakeholder preferences with sensitivity analyses (Reichert et al., 2015). (vi) MCDA is carried out stepwise, thus reducing complexity and increasing transparency. For these reasons, we used MCDA to identify a "Good flood forecast and alert system" for West Africa (see practical research questions RQA, RQB; sect. 1.1). We present the MCDA methods such that they are easily accessible and adaptable to other transdisciplinary projects, e.g., in hydrology, and provide extensive details as blueprint in the Supplementary Information. Based on our experiences, we critically analyzed the MCDA process using the transdisciplinary framework (Table 1) to understand which elements were met by MCDA (RQC, RQD; presented in sects. 4.3 and 5).*~~

## ~~1.3~~2.3 Aims, research gaps, and research questions~~, and structure of paper~~

In this paper we follow two complementary aims: (1) support and document FEWS development for West Africa using a comprehensive MCDA process in close interaction with stakeholders, and including uncertainty; and (2) evaluate participatory MCDA as a transdisciplinary process. To achieve these aims, an inter- and transdisciplinary approach is needed.

~~The aim of the FANFAR project is to co-design and co-develop a pre-operational flood forecasting and alert system at West Africa scale. For the stakeholders in West Africa, and ultimately for thousands of people affected by floods, it is of prime importance to find a good configuration of a flood forecast and alert system. "Good" means that it~~ As outlined in sect. 1.1, there is a lack of FEWS at West Africa scale that sufficiently meet user needs~~satisfies the main needs of those that are meant to use it: the hydrological services and emergency management agencies~~. To interact with stakeholders (from West Africa in the FANFAR case), a transdisciplinary approach is needed. We thus reviewed the transdisciplinary and sustainability science literature to identify frameworks for guiding transdisciplinary projects (sect. 2.1). Despite many shared characteristics, we found differences between frameworks, and only few took the perspective of the natural and earth systems sciences. Moreover, this literature acknowledges a lack of conceptualization and systematic integration of empiric evidence. MCDA can support decisions about suitable FEWS configurations (sect. 2.2). MCDA can integrate different types of data and stakeholder interests, and can deal with uncertainty. However, literature indicates that participatory MCDA integrating stakeholders throughout the process, and focusing on uncertainty, is rarely done in flood risk research. ~~(Research Question A; RQA). Moreover, there was a timing issue. The IT specialists and hydrologists could not wait with technical system development until all stakeholder information had been gathered and analyzed with MCDA. We thus needed to rely on preliminary information early in the project, which was refined as the project progressed (RQB). This leads to following research questions:~~

~~**RQA:** What characterizes a good regional flood forecasting and alert system for West Africa? Is it possible to identify a robust FANFAR system configuration, despite uncertainty (of expert predictions about FANFAR system performance and MCDA model) and possibly different preferences of stakeholders regarding what the system should achieve?~~

~~**RQB:** How can problem structuring at the beginning of the project help focus the development of the FANFAR system to meet the stakeholders' expectations (i.e., before MCDA results are available)?~~

~~A~~Moreover, ~~ccording to~~ the call of this Special Issue on contributions of transdisciplinary approaches to hydrology and water resources management emphasizes: "(…) we need to remind ourselves that a scientific decade on change in hydrology and society requires the perspectives of those disciplines that have traditionally been concerned with society (…). While interdisciplinary conversations have been happening to some extent, transdisciplinary endeavors remain largely undocumented" (Carr et al., 2021). Finally, whether MCDA is suitable to guide transdisciplinary processes in all phases remains an open question. These research gaps lead to following research questions, grouped under the two aims:

~~We aimed to document empirical evidence from the FANFAR project, contributing to knowledge production, learning, and scientific praxis in hydrology. 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). We summarized insights and recommendations of special interest to hydrology praxis when engaging with stakeholders (RQD). The specific research questions are:~~**Aim (1):** Define what constitutes a good FEWS for West Africa using a participatory MCDA process, and document empirical evidence from the FANFAR project, hereby contributing to knowledge production, learning, and scientific praxis in hydrology.

- **RQA:** What characterizes a good regional FEWS for West Africa? Is it possible to identify a robust FEWS configuration, despite uncertainty (of expert predictions about FEWS performance and MCDA model) and possibly different preferences of stakeholders regarding what the FEWS should achieve?

**Aim (2):** Evaluate the suitability of participatory MCDA as a transdisciplinary process.

- **RQB:** What are main characteristics of existing frameworks from transdisciplinary research and sustainability science that are useful for guiding and evaluating collaborative transdisciplinary projects in hydrology research?

- RQC: How suitable is a structured, participatory decision analysis process based on MCDA for guiding large transdisciplinary projects? What worked well or less well in FANFAR? Could interests of a large number of stakeholders be integrated?
- RQD: Is the proposed framework useful for this type of evaluation? What insights and recommendations can we provide for future transdisciplinary projects in hydrology research?

2 The paper is organized as follows: In the Methods, we review main elements of frameworks in the sustainability science and transdisciplinary research literature (sect. 2.1), followed by a framework proposition (sect. 2.1.1). We shortly review literature on MCDA in flood risk management (2.2.1) and motivate using MCDA to structure stakeholder participation (sect. 2.2.2). We present the workshops in West Africa (sect. 2.2.3), and the methods in each MCDA step (sects. 2.2.4 – 2.2.11). In the Results (sect. 3), we present problem structuring steps of MCDA including stakeholder analysis, the final MCDA results, and sensitivity analyses to test robustness of best performing FANFAR system configurations under changing model assumptions. The Discussion follows the research questions: finding a robust FANFAR system configuration with MCDA (RQA, sect. 4.1) and whether early problem structuring helped system design (RQB, sect. 4.2). We analyze MCDA following the proposed framework of transdisciplinary research (RQC, sect. 4.3). We conclude by summarizing insights regarding the framework and recommendations for using an MCDA process in hydrology research, engaging with stakeholders and society (RQD; sect. 5).

# 3 Methods

## 2.13.1 Transdisciplinary Multi-Criteria Decision Analysis (MCDA) process

MCDA in flood risk research

Methodological and epistemological perspectives in transdisciplinary research may still be debated (Jahn et al., 2012), but there needs to be consensus on methods used in a project, and the concept to integrate research (Lang et al., 2012). The FANFAR consortium agreed on Multi-Criteria Decision Analysis (MCDA) as organizing framework to integrate West African stakeholders in the transdisciplinary process and achieve the aim of producing the FANFAR system. MCDA is increasingly popular in hydrology, and specifically related to flood risks. A brief literature search revealed around 50 potentially relevant articles, but only few actually included stakeholders (Web of Science search 25.08.2021; keywords: “MCDA” AND “hydrolog*” AND/OR “flood*”). This corroborates results of two recent reviews (de Brito and Evers, 2016;Abdullah et al., 2021). Both confirm a significant growth in MCDA applications, especially for flood mitigation, while flood preparedness, response, or recovery phases were understudied. Moreover, uncertainty analysis and stakeholder participation were rarely applied, and require future exploration to increase the quality of decision making and success of implementation (de Brito and Evers, 2016). Our search indicates that most papers used MCDA as a technical method to (mathematically) integrate over various indicators, e.g., to support calibration of flood forecasting models (Pang et al., 2019). Recent methodologically interesting papers addressed MCDA coupled with e.g., artificial intelligence (Pham et al., 2021), machine learning (Nachappa et al., 2020), or portfolio decision analysis (Convertino et al., 2019). Combining GIS with MCDA is a trend in hydrology, as in other fields.

Examples include flood risk assessment explicitly including uncertainty (Tang et al., 2018), a method to produce risk maps for a Swiss region (Ronco et al., 2015), and flood risk analyses and mappings for specific cases (e.g., Samanta et al., 2016).

This literature seems to lack integrating stakeholders in the decision analysis process, even in applied cases, a main conclusion of the review by de Brito and Evers (2016). These authors also propose a concept for inter- and transdisciplinary approaches to improve urban resilience in flood risk management using MCDA (Evers et al., 2018), and present a participatory case study in Brazil for flood vulnerability assessment (de Brito et al., 2018). Several studies state that MCDA results are highly susceptible to model assumptions, especially the weights assigned to objectives (de Brito and Evers, 2016). For instance, the sensitivity of MCDA results to weight variability was determined with global sensitivity analysis (Tang et al., 2018). In some of our screened papers, stakeholders actively participated, but again mainly to assign weights. However, most papers we checked seem technical, and stakeholder involvement was usually described in few sentences, presenting the weights in a table, without further elaboration of the process (e.g., Ronco et al., 2015). Concluding our short review, there seems to be a lack of research in hydrology that understands MCDA as a process, and that integrates stakeholders throughout this process.

Overview of transdisciplinary MCDA process

MCDA embraces various methodologies to support complex decisions (e.g., Belton and Stewart, 2002;de Brito and Evers, 2016). We chose MCDA, and specifically Multi-Attribute Value Theory (MAVT; Eisenführ et al., 2010;Keeney, 1982) for several reasons: (i) to develop a complex forecast system, many decisions had to be made. We needed to clarify, which hydrological models and data should be used to produce flood forecasts (e.g., type of meteorological forecast data, streamflow gauge observations), which flood hazard thresholds were appropriate, how to visualize model forecasts, or which distribution channels effectively reach people. MCDA is optimally suited to address such questions. (ii) Because the FANFAR system should be adapted to stakeholder needs, its development relied on close collaboration with nonacademic partners. MCDA is a proven methodology that allows close interaction with stakeholders at various stages of decision making processes, offering a choice of methods for each stage. (iii) We used MAVT because it bases decisions on the objectives that are of fundamental importance to stakeholders. (iv) To evaluate options (i.e., system configurations), MCDA allows comparing and integrating different kinds of scientific and technical data from expert predictions (e.g., accuracy of forecasts, estimated costs of system development) with stakeholders' preferences. Disentangling facts from values can be very helpful, especially if stakeholders have conflicting interests (Gregory et al., 2012a;Keeney, 1982). In most complex decisions, not all objectives can be fully achieved. MCDA explicitly asks stakeholders for the trade-offs they are willing to make. (v) MCDA allows including various types of uncertainty, e.g., of expert predictions with probability theory, or uncertain stakeholder preferences with sensitivity analyses (Reichert et al., 2015). (vi) MCDA is carried out stepwise, thus reducing complexity and increasing transparency. For these reasons, we used MCDA to identify a "Good flood forecast and alert system" for West Africa (see practical research questions RQA, RQB; sect. 1.1). We present the MCDA methods such that they are easily accessible and adaptable to other transdisciplinary projects, e.g., in hydrology, and provide extensive details as blueprint in the Supplementary Information. Based on our experiences, we critically analyzed the MCDA process using the transdisciplinary framework (Table 1) to understand which elements were met by MCDA (RQC, RQD; presented in sects. 4.3 and 5).

A typical MCDA process starts with joint problem framing (Fig. 1Figure 1). For better understanding, we added three main steps of our , as in the co-design step 1 of the transdisciplinary framework to the overview- (see Table 3); problem framing belongs to co-design step 1(Table 1). Not part of MCDA sensu stricto were building the collaborative research team (step 1a), defining research questions, and methodological framework (step 1b), and the boundary object (, namely the FANFAR system FEWS for West Africa; (step 1c;- Table 3Table 1). However, these steps werethis was part of the FANFAR project with joint proposal writing and a kick -off meeting of with European and West African consortium partners (section sect. 3.1.12.2.3). As first step of the MCDA process that includes problem structuring, we undertook did a stakeholder analysis (e.g., Grimble and Wellard, 1997;Lienert et al., 2013;Reed et al., 2009), which is often neglected in MCDA projects. Identifying stakeholders is crucial in any participatory project. This was especially important, as European researchers worked in an unfamiliar African context. Main identified stakeholders that participated in the workshops were representatives from hydrological services, emergency management agencies, river basin organizations, and regional expert agencies. Together with these priority stakeholders, we then identified objectives ("What is of fundamental importance to be achieved by a FANFAR systemFEWS?") and options ("Which FANFAR systemFEWS configurations are potentially suitable to achieve objectives?"). To support these first MCDA steps, diverse "Problem Structuring Methods" (PSMs) are available in other fields of decision analysis (Rosenhead and Mingers, 2001). It is common to combine MCDA with PSMs (reviewed by Marttunen et al., 2017). For a description of sSimilar PSMs as used in FANFAR, we refer to an applicationwere described in a wastewater infrastructure planning example (Lienert et al., 2015). The next steps 5–7 in MCDA (Fig. 1)(Figure 1), belong to the transdisciplinary co-production step 2 (Table 1). Hereby, research is conducted to produces new knowledge, in continuous exchange between scientists from different disciplines and stakeholders. A transdisciplinary process is often iterative (e.g., Jahn et al., 2012;Lang et al., 2012), captured in the FANFAR project with several cycles of workshops with decision makers, end users, and stakeholders ("stakeholders" hereafter) to test, discuss, and improve the FANFAR pre-operational systemFEWS. In the co-dissemination and evaluation step 3, this new knowledge is critically reflected, integrated, and disseminated (Table 1), captured in step 8 of a typical MCDA process (Fig. 1)(Figure 1). After summarizing the workshops in Africa (sect. 3.1.12.2.3), we -focus on the MCDA steps of the MCDA process (sects. 3.1.22.2.4 –3.1.82.2.10).We present MCDA methods such that they are easily adaptable to other transdisciplinary projects, e.g., in hydrology, and provide extensive details as blueprint in the Supplementary Information.

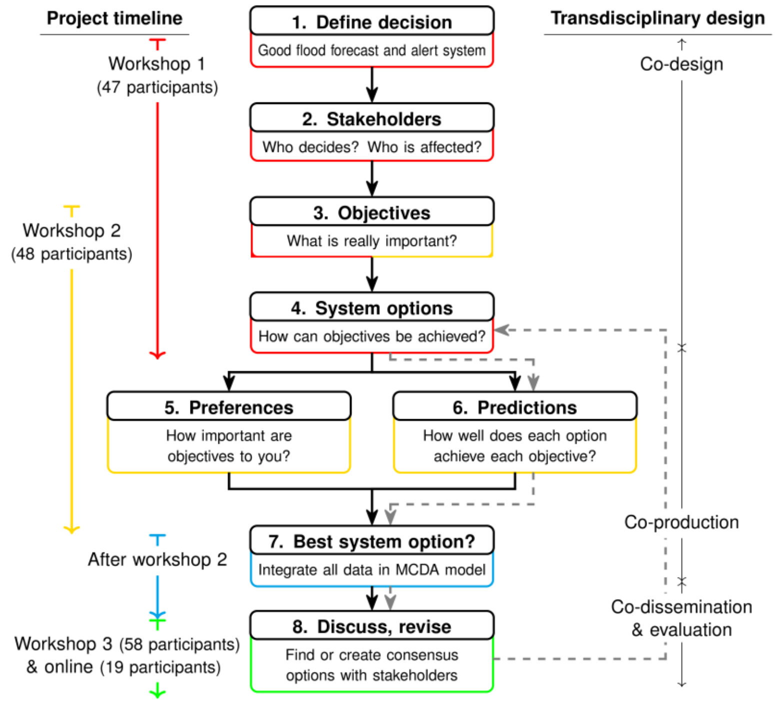



**Figure 1: Multi-Criteria Decision Analysis (MCDA) was ~~is~~ carried out stepwise in the FANFAR project. Explanations see text.**

### ~~2.1.1~~3.1.1 Co-design workshops in West Africa

We carried out three workshops in West Africa, and a FANFAR consortium kick off meeting ~~of the FANFAR consortium (in~~ Norrköping, ~~(~~Sweden, 17–18 January 2018). A fourth workshop was replaced by two half-day online workshops due to COVID-19 (20–21 January 2021), and a final online workshop (~~0~~1 June 2021). The workshops are documented in ~~in a detailed~~ reports ~~(Lienert et al., 2020), and summary reports~~ (FANFAR, 2021;Lienert et al., 2020). At each workshop, West African ~~representatives~~ stakeholders presented the rainy season flood situation in their country ~~during rainy seasons~~ and their experience with the FANFAR ~~FANFAR system~~FEWS. Each workshop hosted extensive technical sessions for experimentation with the latest ~~FANFAR system~~FEWS configuration, and structured technical feedback. Between workshops, the ~~pre-operational system~~ FEWS was adapted to meet requests as well as possible (Andersson et al., 2020a). We also conducted sessions with emergency managers, e.g., about their understanding of flood risk representation to improve FEWS visualizations ~~of the FANFAR system~~ (Kuller et al., 2020). ~~In this paper~~Here, we focus only on interactions at the core of ~~the~~ MCDA ~~process~~.

The first workshop ~~(in~~ Niamey, ~~(~~Niger, 17–20 September 2018) hosted 47 participants from 21 countries, including European and African consortium members ~~from Europe and Africa~~, and representatives from regional and national hydrological service ~~agencies~~ and emergency management agencies ~~on regional and national levels~~ from 17 West and Central African countries ~~in West and Central Africa~~. Main aim was ~~to~~ initiat~~e~~inge the co-design process. For MCDA, we ~~carried out~~used ~~the~~ problem structuring ~~steps~~ (Fig. 1)~~(Figure 1)~~: ~~a~~ stakeholder analysis (sect. 3.1.2~~2.2.4~~); ~~different interactions to~~ identifying fundamentally important objectives of ~~participants~~ stakeholders (~~i.e., what the system should achieve;~~ sect. 3.1.3~~2.2.5~~), and ~~to identify system~~ FEWS configurations that meet ~~the~~ objectives; sect. 3.1.4~~2.2.6~~). The second workshop (~~in~~ Accra, ~~(~~Ghana, 9–12 April 2019) hosted 48 participants from 21 countries. For ~~the~~ MCDA, we consolidated ~~the list of~~ objectives~~,~~ and elicited participants' preferences regarding achieving these objectives ~~in small groups~~ (sect. 3.1.6~~2.2.8~~). Additionally, we collected preference data on the importance of objectives from each ~~individual participant~~ stakeholder with questionnaires. This provided interesting insights into preference formation ~~and changes~~ over time (Kuller et al., in prep.). For the third workshop (~~in~~ Abuja, ~~(~~Nigeria, 10–14 February 2020), ~~the number of~~ participant~~s~~ numbers increased to 58, including representatives from WMO (World Meteorological Organization; https://public.wmo.int/), ECOWAS (Economic Community of West African States; https://www.ecowas.int/), and ~~from~~ 16 West and Central African countries. We discussed main MCDA results. During ~~one of the~~a last online workshop~~s~~, which was attended by 10–19 participants (varying numbers due to internet connection problems), stakeholders completed a survey, providing some feedback for MCDA (sect. 3.1.9).

### ~~2.1.2~~3.1.2 Stakeholder analysis

For the stakeholder analysis (Grimble and Wellard, 1997;Reed et al., 2009), we followed Lienert et al. (2013). ~~The w~~Workshop participants filled in a pen and paper questionnaire in French or English, assisted by two experts. The survey was completed in 2.5 hours by 31 participants in 18 groups (for countries). After receiving information, the participants completed two tables, one for identifying key West African organizations ~~involved in~~that produc~~ing~~e and operat~~ing~~e ~~flood forecast and early warning systems~~FEWS, and one for ~~identifying~~ downstream stakeholders (i.e., "Who might play a role because they use information from such systems in society?"). Each table contained eight tasks: (1) listing key organizations or stakeholders; (2) specifications (e.g., names); (3) their presumed main interests; (4) why they might use ~~the FANFAR system~~FEWS; and (5) appropriate distribution channels. We used a 10 point Likert scale, asking participants to (6) rate the importance of considering each listed stakeholder ~~or organization~~ in the FANFAR co-design process; (7) the presumed influence (power) of each stakeholder for implementing the ~~FANFAR system~~FEWS; and (8) how strongly each ~~stakeholder or organization~~ would be affected by the FEWS performance level~~system (i.e., its level of performance)~~. We cleaned the raw data and categorized stakeholders ~~based on whether~~according to: ~~they are~~ forecast/alert producers or users, ~~their~~ decisional level, sector, and perceived main interest~~. More details see~~ (details see Silva Pinto and Lienert, 2018).

### ~~2.1.3~~3.1.3 Generating objectives and attributes

Generating objectives is key to MCDA (Belton and Stewart, 2002;Eisenführ et al., 2010;Keeney, 1982), since this choice can alter results. Value Focused Thinking guides this step by focusing on what is fundamentally important to stakeholders (Keeney, 1996). However, s~~S~~imply asking is ~~stakeholders is~~ insufficient, and often too few (Bond et al., 2008;Haag et al., 2019c) or too many objectives are produced; we refer to the guidelines in Marttunen et al. (2019). Our stepwise procedure started at the FANFAR kick off meeting in Sweden and continued in the first two West African workshops (details see Lienert et al., 2020). In the first workshop, ~~we split participants into three groups. In the first~~one stakeholder group ~~,~~individual~~s~~ly used an interactive online survey to first brainstorm, then select objectives from a master list (Haag et al., 2019c). Individuals in ~~the~~ a second group used the same procedure ~~in a~~as pen and paper survey, assisted by a moderator. The third group used a means-ends network in a moderated group discussion to find consensus objectives (Eisenführ et al., 2010). Each participant (respectively group)~~,~~ ranked and rated objectives according to importance. Objectives were discussed in the plenary and the most important ones were chosen by majority vote. ~~Between workshops, w~~We post processed objectives to avoid common mistakes such as double counting, ~~and~~ overlaps, or ~~including~~ means objectives (Eisenführ et al., 2010). MCDA objectives are only useful if they discriminate options (~~system~~ FEWS configurations in our case), and we dismissed those not fulfilling this requirement. In the second workshop, we presented a revised list of ~~the~~ 10 most important objectives, including a clear definition of the best and worst possible case for each (see attribute descriptions; ~~for attributes in~~ sect. SI-2.4.1). For instance, the FEWS being available in ~~*for the objective* "~~*2.3 Several languages*~~"~~ ~~, the FANFAR system being available in several languages~~ is the best case, and only in English the worst (~~.~~note: names of objectives, attributes, FEWS configurations, and stakeholder groups are given in Italics; see Fig. 2; Table 1). After discussion, ~~the workshop participants~~stakeholders agreed on the final ~~list of~~ objectives as basis for MCDA. To operationalize objectives, attributes (synonym indicators) are required (Eisenführ et al., 2010). These were developed by experts from the FANFAR consortium. In most cases, we constructed attributes from several sub-attributes (sect. 3.1.5~~2.2.7~~). Sub-attributes or attributes were transformed to a value ~~with help of~~using marginal value functions (sect. 3.1.6~~2.2.8~~)~~. They were aggregated to a single total value with the MCDA model (sect. 2.2.9)~~.

### ~~2.1.4~~3.1.4 Generating ~~system~~ FEWS configurations

Different plausible ~~FANFAR system~~FEWS configurations were generated in the first workshop, in three moderated group sessions. Two groups used the "Strategy Generation Table" (Gregory et al., 2012b;Howard, 1988), and one "Brainwriting 635" (Paulus and Yang, 2000) combined with "Cadavre Exquis" (write words on ~~a~~ paper, fold, ~~and~~ give to next person). The Strategy Generation Table ~~is a systematic procedure that~~ allowed pre-structuring FEWS elements ~~of the FANFAR system~~ (e.g., observed variables, ~~models for~~ forecast production models, language). ~~The s~~Stakeholders chose elements forming suitable ~~system~~ FEWS configurations ~~("strategies")~~ with help of questions ~~such as~~: "The most easy to use ~~system~~FEWS", or the "Most robust FEWS~~system~~ working well given West African boundary conditions ~~in West Africa~~ (e.g., internet or power supply problems)". Brainwriting 635 allowed for ~~open,~~ interactive brainstorming, using the same ~~strategies~~ questions~~as the other groups~~. ~~All~~ We discussed all ~~FANFAR system~~FEWS configurations ~~were discussed~~ in the plenary. As part of post processing, FANFAR consortium members ~~additional~~ created technically interesting ~~system~~ FEWS configurations ~~were created by FANFAR consortium members~~. ~~For readers unfamiliar with the methods, w~~We provide details in the Supplementary Information for readers unfamiliar with the methods (sect. SI-1.1).

### ~~2.1.5~~3.1.5 Predicting performance of each ~~system~~ FEWS configuration

Part of the MCDA input data are scientific predictions (Fig. 1)~~(Figure 1)~~, based on estimates or models of the performance level ~~of achievement~~ for each objective (Eisenführ et al., 2010). We used expert estimates (O'Hagan, 2019) by interviewing

FANFAR consortium members in July–August 2019. First, experts developed attributes (sect. 3.1.32.2.5), in most casesmostly constructed from sub-attributes. They estimated the outcome of each FANFAR systemFEWS configuration for each (sub-) attribute, i.e., the likely level of each attribute (e.g., likely operation costs) and gave uncertainty ranges regarding their predictions. For constructed attributes, we integrated the predictions of the sub-attributes into one final value using a weighted sum (weights defined by experts; sect. 3.1.62.2.8). We aggregated the uncertainty of each sub-attribute into a single uncertainty distribution with 1’000 Monte Carlo simulations. To characterize the resulting aggregated uncertainty, we used a normal distribution with mean (of Monte Carlo simulation), and standard deviation (¼ of the 95 % confidence interval from simulation) as input in the MCDA (sect. 3.1.72.2.9).

Example: The oObjective “*1.1 High accuracy of information*” consists of three sub-attributes: KGE index for 1, 3, and 10 day forecasts (Kling-Gupta Efficiency; Gupta et al., 2009). The KGE is one possible accuracy index for hydrological model evaluation, e.g., to estimate the error of predicted vs. observed values. For each FANFAR systemFEWS configuration and lead day, the expert estimated the expected KGE. The KGE index number was transformed to a value, ranging from 0 (worst) to 1 (best), with a nonlinear marginal value function, elicited from the expert. We aggregated the lead day values into a single value [0:1] with a weighted sum, where the accuracy of the 1 day forecast received a weight of 0.5, the 3 day forecast of 0.4, and 10 day forecast of 0.1. Details for predicting system performance (expected achievement of stakeholderattribute level objectives) see sect. SI-2.4.

### 2.1.63.1.6 Eliciting stakeholder (or expert) preferences

***Marginal value functions.*** Subjective preferences of stakeholders enter the MCDA model on equal footing to expert predictions (Fig. 1)(Figure 1). Preference elicitation is an important, sensitive step during which many biases can occur (Montibeller and von Winterfeldt, 2015). It is crucial to follow recommendations (Eisenführ et al., 2010). Marginal value functions convert the attribute levels for each objective (e.g., KGE index for “*1.1 High accuracy of information*”) to a common scale ranging from 0 (worst possible achievement of this objective) to 1 (best achievement). This allows integrating different attributes with various different units into one model, e.g., the KGE index with operation costs (€ / y$^{-1}$ear), and development time (days). As default, a linear marginal value function can be used. However, nonlinear value functions usually better capture preferences. In FANFAR, most attributes are relatively technical, requiring expert knowledge. We therefore thus elicited shapes of value functions from experts (sect. 3.1.52.2.7; details, including figures of value functions, see sect. SI-2.4.1). For each sub-attribute, we mostly created seven evenly spaced levels (worst, very bad, bad, neutral, good, very good, and best). Experts then assigned attribute numbers (e.g., KGE index for 3 day forecasts) to each level based on their experience. We transformed attribute levels to [0:1] values using linear interpolation between levels. As example, the KGE index ranges from minus infinity (worst case, value 0) to 1 (best case, value 1; Table SI-8). For each attributes consisting of sub-attributes, we elicited a nonlinear marginal value function for each sub-attribute (FigureFig. SI-5), allowing aggregation into one onesingle value. Because we already used elicited nonlinear value functions to construct the composite attribute, we used a linear value function for these in ~~the~~ MCDA (sect. 3.1.7~~2.2.9~~).

***Weights.*** In the second FANFAR workshop, we elicited ~~the~~ weights ~~by dividing participants into~~from five groups~~,~~ ~~-~~according to language (French F, English E) and professional background (Emergency m~~M~~anagers ~~EM~~, Hydrologists ~~HY~~). The two French speaking groups used the ~~the~~ Swing method (Eisenführ et al., 2010): eight emergency managers (group *1. Emergency-F*~~ID: G1A_EM_F, where G1A = group~~), and 11 hydrologists (two sub-groups~~,~~ *2A. and 2B. Hydrology-F*~~G2A_HY_F, G2B_HY_F~~). The two English speaking groups used an adapt~~ed~~ation ~~of~~ Simos' revised card procedure (Figueira and Roy, 2002;Pictet and Bollinger, 2008), hereafter Simos card: 14 hydrologists (*3. Hydrology-E*~~G3A_HY_E~~), and three emergency managers (*4. Emergency-E*~~G4A_EM_E~~). We ~~separately~~ elicited weights from three AGRHYMET experts with Simos card ~~method~~ (*5. AGRHYMET-E*~~G5A_AGRHYMET~~). Stakeholders can be uncertain about their preferences, or groups may disagree. For Swing, we avoided forcing participants to reach group consensus and encouraged ~~discussing~~ discussion of diverging opinions,~~.~~ ~~This~~ result~~ing~~ed in a range of ~~the stakeholders'~~ weights ~~preferences~~. We took the mean as main weight and considered strong deviations (difference in weights > 0.2 compared to mean) in ~~later~~ sensitivity analyses (sect. 3.1.8~~2.2.10~~). For Simos' card ~~procedure~~, two additional weight sets ~~were used,~~ result~~ed~~ing from eliciting a range for one variable. The moderator recorded important comments to inform ~~the~~ sensitivity analyses (Table SI-3). ~~In the group of~~For French speaking hydrologists, two diverging preference sets emerged from the start, which we analyzed separately (~~*G*~~*2A*, ~~*G*~~*2B*). For interested readers, we give details of ~~these~~ standard MCDA weight elicitation ~~procedures~~ (sect. SI-1.2). To check for the validity of the additive aggregation model (sect. 3.1.7~~2.2.9~~), we shortly discussed implications in the weight ~~elicitation~~ sessions using elicitation procedures from our earlier work (Haag et al., 2019a;Zheng et al., 2016).

### ~~2.1.7~~3.1.7 MCDA model integrating predictions and preferences

The MCDA model integrates expert predictions with stakeholder preferences, and calculates~~.~~ ~~An aggregation model is used to calculate~~ the total value of each ~~alternative~~ FEWS configuration (= alternatives; Eisenführ et al., 2010). A finite set of FEWS alternatives $A = \{a, b, \ldots\}$ are evaluated regarding the predicted outcomes on every objective (respectively attribute). We denote ~~the~~ predicted outcomes (sect. 3.1.5~~2.2.7~~) as $x_a = (x_{a,1}, \ldots, x_{a,n})$, with $x_{a,i}$ ~~being~~ the level of an attribute $i$ that measures a predicted consequence of ~~system~~ FEWS ~~configuration~~ $a$ *(or b, c, …)*. The total value $v(x_a)$ of ~~system configuration~~FEWS $a$ is calculated with a multi-attribute value function, $v(x_{a,1}, \ldots, x_{a,n}, \Theta)$. The resulting total value $v(x_a)$ of each FEWS~~system configuration lies~~is between 0 (all objectives achieve ~~only the~~ worst level) and 1 (all objectives ~~are on the~~achieve best ~~attribute~~ level ~~that can be achieved~~ given the ~~defined~~ attribute ranges). A rational decision maker ~~would~~ chooses the ~~configuration~~ FEWS with the highest value. ~~Most c~~Commonly, an additive model ~~based on marginal value functions~~ is used~~, as in Eq. (1), but non-additive models as in Eq. (3) are also possible~~:

$$\boldsymbol{v(x_1, x_2, \ldots, x_n, \Theta) = \sum_{i=1}^{n} w_i \cdot v_i\,(x_i, \theta)} \quad \text{(Eq. 1)}$$

with parameters $\Theta = (w_1, \ldots, w_n, \theta)$, where $w_i$ is the weight of attribute $i$, with $0 \leq w_i \leq 1$, and $\sum_{i=1}^{n} \boldsymbol{w_i} = \mathbf{1}$, (Eq. 2)

and where $v_i(x_i, \theta)$ is the value for the predicted consequence $x_i$ of attribute $i$ of system configurationFEWS $a$. This value is inferred with help of the marginal value function (sect. 3.1.62.2.8).

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., 2019a;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, using some examples, whether they agree with objectives being preferentially independent, and as consequence with the full compensatory effect. In all five groups this was not the case. We therefore used a non-additive model with less strict requirements, the weighted power mean with an additional parameter $\gamma$ that determines the degree of non-compensation, see Eq. (3):

$$\boldsymbol{v}(\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_n}, \boldsymbol{\Theta}) = \left(\sum_{i=1}^{n} \boldsymbol{w_i} \cdot \boldsymbol{v_i}(\boldsymbol{x_i}, \boldsymbol{\theta})^{\gamma}\right)^{\mathbf{1}/\gamma} \quad \text{(Eq. 3)}$$

If $\gamma = 1$, we are back to the additive model in Eq. (1). We used a value for $\gamma = 0.2$, based on stakeholder input from stakeholders (sect. 3.1.62.2.8), which is closer to a weighted geometric mean ($\gamma \rightarrow 0$). We shortly explain and visualize implications of the power mean in sect. SI-1.3 . For further details we refer to (details see Haag et al., 2019b).

We calculated MCDA results in our new open source software “ValueDecisions” (Haag et al., subm.), . ValueDecisions is based on the software and programming language R (R Core Team, 2018), earlier R scripts developed in our group (e.g., Haag et al., 2019b), and R “utility” package (Reichert et al., 2013). R scripts were rendered as web application for ValueDecisions with the “shiny” package (Shiny, 2020). We used R for aAdditional analyses were implemented directly in R: aggregating uncertainty of sub-attributes, weight visualization, and statistical analysis of sensitivity analyses.

#### 2.1.83.1.8 Uncertainty of predictions and preferences

***Uncertainty of predictions:*** Probability theory is used in MAVT (Reichert et al., 2015). We defined uncertainty distributions from expert predictions for each attribute (sect. 3.1.52.2.7). We calculated aggregated values of each system FEWS configuration across all objectives (sect. 3.1.72.2.9), drawing randomly from the attributes’ uncertainty distributions in 1’000 Monte Carlo simulation runs. We analyzed rank frequencies: how many times in 1’000 runs each FANFAR systemFEWS configuration achieved each rank.

***Sensitivity analyses of aggregation model and stakeholder preferencesweights:*** Local sensitivity analyses are commonly used to check the sensitivity of MCDA results to diverging preferences (e.g., Eisenführ et al., 2010;Zheng et al., 2016;Haag et al., subm.). We checked the sensitivity of MCDA results to other aggregation models and changed weights. We used setting S0 as default, , and comparinged it with results with those of a separate MCDA for each new setting with changed preference input parameters (settings are summarized in results the Results, Table 2Table 3; details see sect. SI-1.4). For each setting, we compared mean ranks of ~~FANFAR system~~FEWS configurations from 1’000 Monte Carlo simulation runs with ~~those of~~ the default MCDA (S0). We used~~ing~~ ~~the~~ nonparametric Kendall’s τ correlation coefficient (Kendall, 1938) to measure rank reversals (as in Zheng et al., 2016). To test ~~implications of~~ the aggregation model (sect. 3.1.7~~2.2.9~~), we recalculated the MCDA for other reasonable models (Haag et al., 2019a; settings S11–S14; Table 2). For weights, ~~the weight of~~ we changed the weight of one objective, ~~is changed,~~ while ~~the~~ ratios of all others w~~weights e~~are kept constant~~,~~ and renormalized ~~so that the sum of weights remains 1~~. For ~~a thorough~~more method explanations ~~we refer to~~see Eisenführ et al. (2010); ~~and give some insight~~details for readers not familiar with MCDA see ~~in~~ sect. SI-1.4. Consistency checks during weight elicitation with ~~the French speaking emergency managers (~~group *1. Emergency-F*~~G1A)~~ revealed an inconsistency, ~~which resulted~~and ~~in~~ strongly different weights (~~Figure~~Fig. SI-3). We tested ~~their effect~~it ~~on results~~ in sensitivity analysis S21 (Table 2~~Table 3~~). For Swing ~~weight~~ weights~~elicitation~~, ~~workshop participants~~stakeholders stated ranges, which we ~~. We~~ tested ~~the sensitivity of results~~ if the difference between the ~~the~~ maximum or minimum ~~weight~~ from the ~~the~~ average weight exceeded Δ = 0.02 (S22). ~~Similarly, f~~For Simos’ card~~ method~~, we tested alternative weight sets resulting from ranges (S23). It is common to test ~~other~~ interesting objectives by doubling the elicited weight. We did this for ~~“~~*2.3 Several languages*~~”~~, because its importance might have been underestimated (S31).

***Cost-benefit visualizations*** are an additional way to check the robustness of results (e.g., Liu et al., 2019). ~~This visual analysis is based on the resulting MCDA, including stakeholders’ preferences.~~ We ~~did this,~~ used~~ing~~ standard setting S0 without prediction uncertainty (Table 2~~Table 3~~) for this visual analysis. For reasons of space, we refer to sect. SI-2.9.

### ~~2.1.9~~3.1.9 Discuss results with stakeholders, feedback

We ~~presented and~~ discussed ~~preliminary~~ first MCDA results ~~with stakeholders~~ in the third stakeholder workshop. ~~Because t~~W~~he fourth w~~orkshop four ~~had to be~~ was carried out online due to COVID-19, and we were not able to thoroughly discuss results. ~~However, w~~We did assess~~ed~~ stakeholder perceived satisfaction with ~~FANFAR system~~FEWS performance during the 2020 rainy season with an online survey, asking ~~to answer~~ following questions for each objective: (a) How much does ~~FANFAR~~ the FANFAR FEWS currently fulfill this objective? (b) Would you use the ~~FANFAR system~~FEWS in ~~the~~ future if it remains as is? (c) What is the minimum acceptable to you? This means: below which level would you NOT use the ~~FANFAR system~~FEWS? (details see sect. SI-1.5). ~~The survey was filled out by 12 participants, resulting for our 10 objectives in 10 x 12 = 120 responses to each question.~~

## ~~2.2~~3.2 ~~Evaluation~~ Conceptual framework for transdisciplinary ~~MCDA~~ process

~~We use common elements~~ Based on the reviewed sustainability science and transdisciplinary research literature (sect. 2.1), we identified common elements. We aimed to set up a useful framework that allows evaluating transdisciplinary processes in hydrology research, including the MCDA in FANFAR. ~~from these frameworks to analyze the transdisciplinary co-design process based on MCDA.~~ We ~~emphasize~~ included elements that are crucial for ~~the~~ practice-oriented ~~FANFAR~~ projects such as FANFAR, and are , aiming to produce a user friendly, but technically complex flood forecast and alert system. Some elements are especially relevant to the earth systems sciences, e.g., climate assessment (Lemos and Morehouse, 2005). , but received less attention in the social science oriented sustainability and transdisciplinary research literature: explicit consideration of uncertainty, and the interdisciplinary effort needed for tackling technically complex problems (Lemos and Morehouse, 2005;Mauser et al., 2013). We follow a stepwise timeline, as proposed by many (Jahn et al., 2012;Lang et al., 2012;Mauser et al., 2013), using the terminology by Mauser et al. (2013): (i) co-design; (ii) co-production, and (iii) co-dissemination of knowledge (to which we add evaluation), involving academia and stakeholders throughout. For each phase, we include specific elements appropriate for our context (Table 1).

## 4 Results

HereThe , we focus on the MCDA results are, ordered as in the the Methods section (sect. 3.12.2) for . This provides answersanswering to the practical research question (RQA ()sect. 4.1–4.8). We present the evaluation framework in sect. 4.9 (RQB). Based on these MCDA results, and the framework, we address the other research questionsRQC in the Discussion.

### 2.34.1 Stakeholder analysis

Of 249 stakeholders listed by workshop participants, 68 distinct types remained after data cleaning (details see Silva Pinto and Lienert, 2018). Stakeholders that were perceived to have high influence and also potentially being highly affected by the FANFAR FANFAR systemFEWS were national entities for disaster management, water resources, and infrastructure, who were already well represented in the co-design processFANFAR (details Table SI-4). Several sSpecific organizations were also perceived as highly important and affected, such ase.g., the “Autorité du Bassin de la Volta” (ABV), who participated in workshops, and the consortium member AGRHYMET, an organization representing 13 West African states and consortium member. Other important/affected parties were mainly stakeholders receiving flood forecasts and alerts such as NGO’s, electricity utilities, dam managers, and the agricultural sector. The Red Cross and environmental protection agencies were perceived to have slightly lower importance/affectedness, among others. Civil society (e.g., communities) would be strongly affected, but have limited decisional influence on developing the FANFAR systemFEWS. In contrast, the media, industry, and commerce were perceived to have more influence, but would not be strongly affected. Such outlier stakeholders could potentially provide a different view to the FANFAR systemFEWS.

### 2.44.2 Objectives and attributes

The oObjectives covered a broad range of issues that of fundamental importance to the stakeholders considered as fundamentally important in view of a *Good flood forecast and alert systemFEWS* for West Africa (objectives, attributes, and system configurations are given in Italics hereafter; Fig. 2Figure 2). Several oSome objectives concerned quality requirements, for the forecasted flood information, grouped under as *1. High information accuracy and clarity*,, and aspects of *2. Good information* *access* ~~by users~~ such as accounting for language diversity. Aspects of *3. Low costs* and ~~longer term~~ *4. High sustainability* were also important, e.g., ~~such as~~ *42. Skilled~~ful human resources~~labor* ~~available~~ in West Africa, capable of maintaining, operating, and accessing the ~~system~~FEWS. Each objective is characterized by an attribute, for operationalizing the objectives' achievement ~~of objectives~~ (Fig. 2~~Figure 2;). Details of~~ attribute calculations ~~are given in~~see sect. SI-2.4).

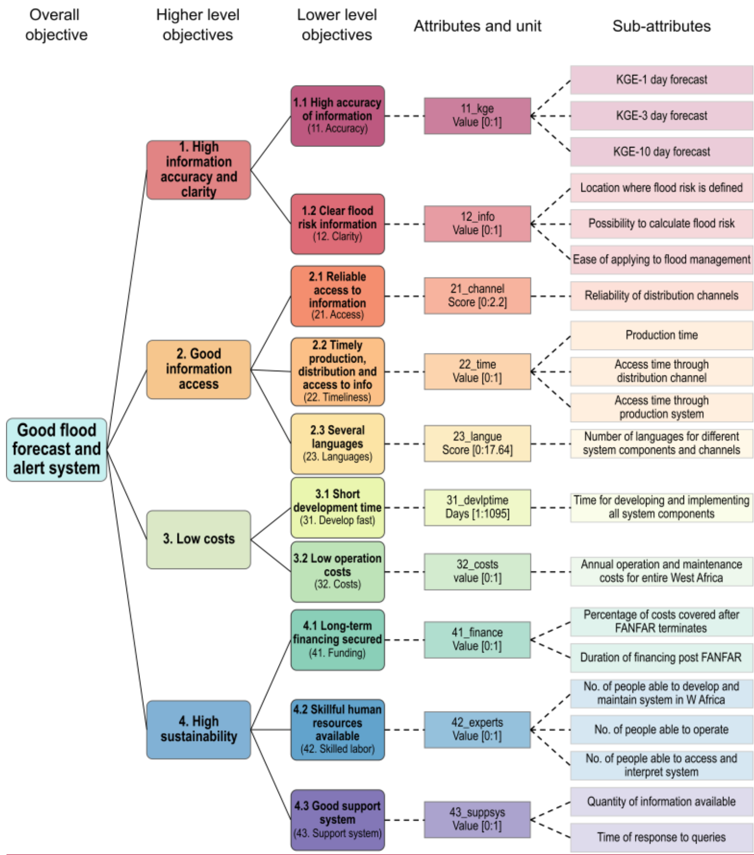



**Figure 2. Objectives hierarchy. From left to right: overall objective, four higher level fundamental objectives, 10 lower level fundamental objectives (short names in brackets) and corresponding attributes, attributes' unit (usually a value) and range [square brackets], from worst (usually value = 0) to best (usually value = 1). Most attributes were constructed from ~~several~~ sub-attributes (far right).**

### ~~2.5~~4.3 ~~FANFAR system~~FEWS configurations

Stakeholders generated six ~~system~~ FEWS configurations in workshop sessions (b to g; Table 1~~Table 2~~)~~ in workshop sessions~~. Experts of the FANFAR consortium developed five configurations (h to k) to cover important technical aspects such as using ~~more~~ refined hydrological models, e.g., redelineation and recalibration of the World Wide HYPE model to West Africa (Andersson et al., 2020b), and including earth observations (EO) from satellites. ~~Configurations~~ FEWS were constructed in separate sessions with experts from AGRHYMET for the forecast production system, and with stakeholders for the user interface IVP (Interactive Visualization Portal). They were combined to form plausible combinations of various ~~FANFAR system~~FEWS elements (summary of important features in Table 1~~Table 2~~; ~~for~~ all ~~system~~ FEWS elements see Tables SI-6 and ~~Table~~ SI-7). Configuration ~~*a_Fast-dev*~~*a. Status quo* represents roughly the state of the ~~first~~initial FEWS version ~~of the FANFAR system~~, when stakeholders started experimentation and giving feedback in the first workshop.

**Table 1~~2~~. Overview of 11 ~~FANFAR system~~FEWS configurations. Selected main characteristics: recent hydrological observation data types (HydObs; WL: water level, Q: river discharge, EO: Earth Observations) & meteorological input/forcing data (MetF; HydroGFD; HydroGFD3 (Berg et al., 2020; improved version); HydroGFD-WA: HydroGFD2 adjusted by West African meteorological observations; Am: American meteorological forecasts (e.g., GFS); Ens: ECMWF ensemble meteorological forecasts); hydrological models (WWH: World-Wide HYPE); forecast output variables (Q: river discharge; WL: water level, P: precipitation; E: evaporation; SM: soil moisture, WQ: water quality); data download (Excel: table for selected station); distribution channels (Web: web visualization; H-TEP: login to H-TEP to download data; FTP: FANFAR and national FTP; API: Application Programming Interface; SoMed: Social Media e.g., WhatsApp; ConMed: conventional media e.g., radio, TV; Tradit: traditional word of mouth) & automatization (Automatic: automatic push of data to distribution channels; Manual: automatic processing with manual control of distribution by operator); flood hazard reference threshold types (RP Sim: return period based on simulations; RP Obs: return periods based on observations at gauged locations; HistY: selected historic year; Local: user defined thresholds for specific location); language of user interface (En: English; Fr: French; Pt: Portuguese; Ar: Arabic).**

| **Name~~ID~~** | **Configuration** | **Hydrological observations & meteorological forcing** | **Hydrological models** | **Forecast output variables** | **Data download** | **Distribution channels & automatization** | **Flood hazard thresholds** | **Language** |
|---|---|---|---|---|---|---|---|---|
| a~~ Fast dev~~. Status quo | Least resources for development: no new features, status quo | HydObs: none; MetF: HydroGFD2 | Niger HYPE | Q | None | Web; Automatic | RP Sim | En |
| b. Resource friendly ~~Res user~~ | Least resources for users (e.g., skilled personnel, stable internet and power) | HydObs: in situ WL, Q; MetF: HydroGFD3 | WWH | Q, WL, P, E, SM | Excel, maps, graphs | Web, H-TEP, SMS, Email, SoMed, ConMed, Tradit; Manual | RP Sim, RP Obs, HistY, Local | En, Fr, Pt, Ar |
| c. User friendly ~~Easy use~~ | Most easy to use for producing and interpreting forecasts and alerts | HydObs: EO WL; MetF: HydroGFD2 | Niger HYPE | Q, WL, P, E | Excel, graphs | Web, SMS, SoMed, ConMed, Tradit; Automatic | RP Sim, HistY | En, Fr, Pt |
| d. Fast alerts ~~Fast~~ | Fastest system for producing and distributing forecasts and alerts | HydObs: EO WL; MetF: HydroGFD2 | Niger HYPE | Q | None | Web, SMS, Email, SoMed, ConMed, Tradit; Automatic | RP Sim | En |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **e. Consensus ~~Consent~~** | Highest consensus: system elements that West African stakeholders mostly agreed on | HydObs: in situ WL, Q, EO WL; MetF: HydroGFD-WA, Am, Ens | Niger HYPE, WWH | Q, WL, P, E, SM | Excel, maps, graphs | Web, H-TEP, SMS, Email, SoMed, ConMed, Tradit; Manual | RP Sim, HistY | En, Fr, Pt |
| **f. ~~–~~Robust** | Most robust in West Africa: works despite problems in e.g., data collection | HydObs: EO WL; MetF: HydroGFD2 | Niger HYPE, WWH | Q, WL, P, E, SM | Excel, maps, graphs | Web, H-TEP, SMS, Email, SoMed, ConMed, Tradit; Manual | RP Sim, RP Obs, HistY, Local | En, Fr, Pt, Ar |
| **g. ~~–~~Attractive** | Most attractive to West African stakeholders: includes many desired features, similar to *h. Fully equipped*, but simpler distribution~~Most attractive in West Africa: many desired features, similar to *h_Equipp*, but simpler distribution~~ | HydObs: in situ WL, Q, EO WL; MetF: HydroGFD-WA, Am, Ens | Niger HYPE, WWH | Q, WL, P, E, SM, WQ | Excel, maps, graphs | Web, H-TEP, SMS, Email, SoMed, ConMed, Tradit; Manual | RP Sim, RP Obs, HistY, Local | En, Fr, Pt, Ar |
| **h. Fully ~~–eE~~quipped** | Fully equipped: all system elements, except recalibrated HYPE models | HydObs: in situ WL, Q, EO WL; MetF: HydroGFD-WA, Am, Ens | Niger HYPE, WWH | Q, WL, P, E, SM, WQ | Excel, maps, graphs | Web, H-TEP, FTP, API, SMS, Email, SoMed, ConMed, Tradit; choice (Automatic or Manual) | RP Sim, RP Obs, HistY, Local | En, Fr, Pt, Ar |
| **i. ~~–~~Calibrated** | Recalibrated HYPE models | HydObs: none; MetF: HydroGFD2 | Recalibrated WWH | Q, WL, P, E, SM | Excel, maps, graphs | Web, H-TEP, SMS, Email, SoMed, ConMed, Tradit; Manual | RP Sim | En, Fr, Pt |
| **j. ~~–~~Calibrated + ~~-~~EO** | Recalibrated HYPE models and EO data | HydObs: EO WL; MetF: HydroGFD2 | Recalibrated WWH | Q, WL, P, E, SM | Excel, maps, graphs | Web, H-TEP, SMS, Email, SoMed, ConMed, Tradit; Manual | RP Sim | En, Fr, Pt |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **k. –Cali-brated + -EO + In-situ** | Recalibrated HYPE models and EO data and in situ data | HydObs: in situ WL, Q, EO WL; MetF: HydroGFD2 | Recali-brated WWH | Q, WL, P, E, SM | Excel, maps, graphs | Web, H-TEP, SMS, Email, SoMed, ConMed, Tradit; Manual | RP Sim | En, Fr, Pt |

### ~~2.6~~4.4 Predicted performance of each ~~system~~ FEWS configuration

Based on expert predictions but excluding stakeholder preferences, no ~~FANFAR system~~FEWS configuration achieved ~~the~~ best levels of all objectives (Fig. 3~~Figure 3~~; details see sect. SI-2.4, raw input data for MCDA modelling Table SI-30). This illustrates the impossibility to design a perfect ~~system~~FEWS, given the inherent trade-offs between achieving objectives. For instance, the status quo pre-operational ~~system configuration~~FEWS ~~*a Fast dev*~~*a. Status quo* achieved the highest values for objective *31. Develop fast (short development time)* ~~*short development time*~~, and *32. C~~c~~osts*, but scored low on many others such as *11. Accuracy,* ~~*accurate,*~~ *12 Clarity, 21. Access, and 22.* ~~*clear, reliable, and*~~ *T~~t~~imeli~~y~~ness* of information. ~~Configurations~~ FEWS achieving high levels for objectives of *1. High information accuracy and clarity*~~,~~ ~~inevitably need longer~~ cannot well achieve *31.* ~~*development*~~ *Develop fast*~~*time*~~ ~~and~~ at low ~~*have higher*~~*32. C~~c~~osts*. Therefore, it is not possible to clearly determine the "best" ~~configuration~~ FEWS based on only the predicted performance (Fig. 3~~Figure 3~~). We require ~~input from~~ stakeholder inputs ~~concerning~~ about ~~which~~ the importance of objectives ~~are most important~~ (sect. 4.5~~3.5~~, sect. 4.6~~3.6~~).

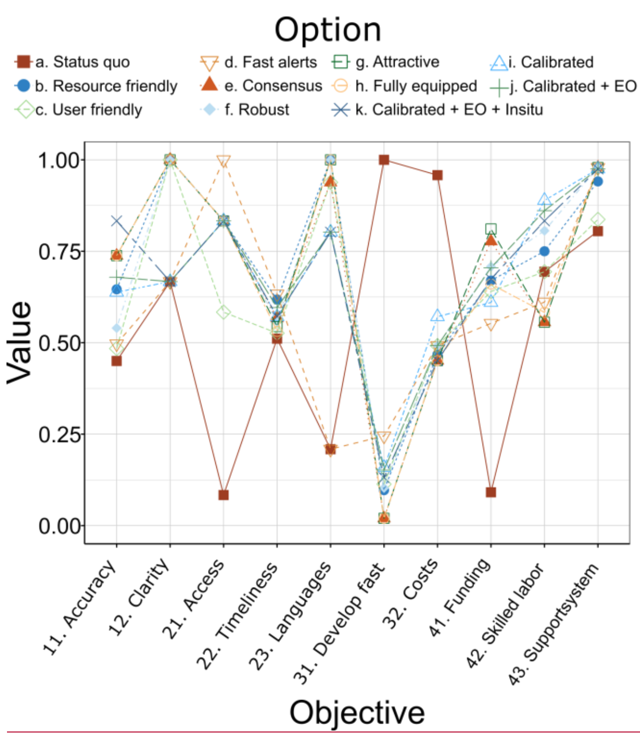
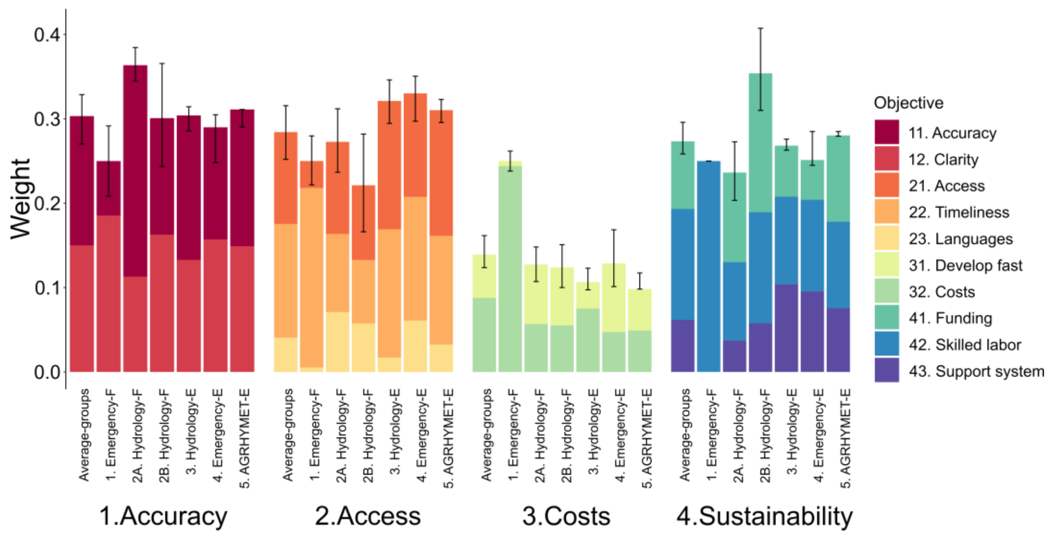



**Figure 3. Predicted value (y-axis) of 11 ~~FANFAR system~~FEWS configurations (options a–k; symbols) for 10 objectives (x-axis), based on expert predictions, but not including stakeholder preferences. Value 1: this ~~system~~ FEWS configuration achieved the best level of this objective; 0: ~~configuration~~ FEWS achieved ~~the~~ worst level, given the ranges of ~~the~~ underlying attributes (i.e., it is a relative scaling from best to worst).**

### ~~2.7~~4.5 Stakeholder preferences

The elicited weights (*w*) for the four higher level objectives were similar for all groups (*w* = total bar length; Fig. 4~~Figure 4~~), except for ~~those of~~ the French speaking emergency managers (~~*G1A*~~*1. Emergency-F*). These gave a high weight (*w* = 0.25) to *3. Low costs*, which was least important for the others (0.1–0.12). ~~G1A~~ They reasoned that all four higher level objectives are equally important in emergency situations with a connected chain of events. In contrast, the higher level objectives *1. High information accuracy and clarity*, and *2. Good information access* were ~~generally~~ usually ~~regarded as~~ most important ~~by~~ for the other groups. There were some notable differences in importance of lower level objectives. Again, group *1. Emergency-F* ~~the French speaking emergency managers (G1A)~~ was~~ere~~ 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~~FEWS development should focus on achieving fast access to flood alerts (*22. ~~–timely–info~~Timeliness*; w = 0.21) and on personnel that can deal with this information (*42. Skilled labor~~–human–resour~~*; w = 0.25). ~~The w~~Weights~~ sets~~ in the other groups were ~~overall~~ more balanced (details ~~see~~ sect. SI-2.6). There ~~were~~ was ~~different levels of consensus~~varying agreement about weights within a group, reflected in the length of ~~the~~ error bars (Fig. 4~~Figure 4~~).




**~~Figure~~Fig. 4. Weights (y-axis) assigned to higher level objectives (blocks, *1. A~~–a~~ccuracy*, *2. A~~–a~~ccess*, etc.) colored by weights of lower level objectives (*11. ~~–accur–info~~Accuracy*, *12. Clarity~~–clear–info~~*, etc.), averaged over all six stakeholder groups (*Average-groups~~–All–SH~~*), and for each group (*1. Emergency-F~~G1A–EM–F~~*, *2A. Hydrology-F~~G2A–HY–F~~*, etc; x-axis). Error bars: uncertainty of elicited preferences~~ statements~~, i.e., the sum of uncertainties of all lower level objectives within the branch of the respective higher level objective. Per definition all weights of a group sum up to 1.**

### ~~2.8~~4.6 MCDA model results

No ~~FANFAR system~~FEWS configuration clearly outperformed the others for all stakeholder groups in the standard MCDA (setting S0; Table 2~~Table 3~~) that did not consider uncertainty (Fig. 5~~Figure 5~~; details see Table SI-32; Table SI-33). The ~~early stage configuration~~FEWS at the beginning of the project (*a. Status quo~~–Fast-dev~~*) achieved lowest total values (*v* < 0.46) and ~~the~~ last ranks for all stakeholder groups, except group *1. Emergency-F* ~~the French speaking emergency managers~~ (~~group G1A,~~

*v* = 0.64, rank 5). This ~~can be attributed~~is caused by ~~to~~ their different weight preferences~~ of G1A~~. All other ~~configurations~~ FEWS generally reached high values for all groups, with ~~only~~ small differences between groups. The total value ranged from *v* = 0.55 in the worst case (*d.* ~~–~~*Fast alerts* for group *2A. Hydrology-F*~~G2A~~) to 0.70 (*b. Resource friendly*~~*b_Res-user,*~~ for *3. Hydrology-E*~~G3A~~). ~~Indeed, t~~This ~~configuration~~ FEWS *b. Resource friendly*~~*b_Res-user*~~ seemed somewhat better than the others, achieving a high value for all groups (*v* = 0.65–0.70)~~ for all groups~~, thus reaching the first rank for all, again with exception of group *1. Emergency-F*~~G1A~~, for which it still achieved the second rank. For better understanding [0,1] values can be interpreted as percentages, and *b. Resource friendly*~~*b_Res-user*~~ ~~was able to~~ achieved 65–70~~-~~% of the ideal case over all objectives in all stakeholder groups. FEWS c~~C~~onfigurations *f.* ~~–~~*Robust*, *i.* ~~–~~*Calibrated*, *j.* ~~–~~*Calibrated + EO*~~-EO~~, and *k. Calibrated + EO + Insitu*~~*k_Cal-EO-situ*~~ also performed well (0.63–0.70) for all groups, while *c. User friendly*~~*c_Easy-use*~~, and *d. Fast alerts*~~*d_Fast*~~ achieved the lowest values (0.55–0.64).

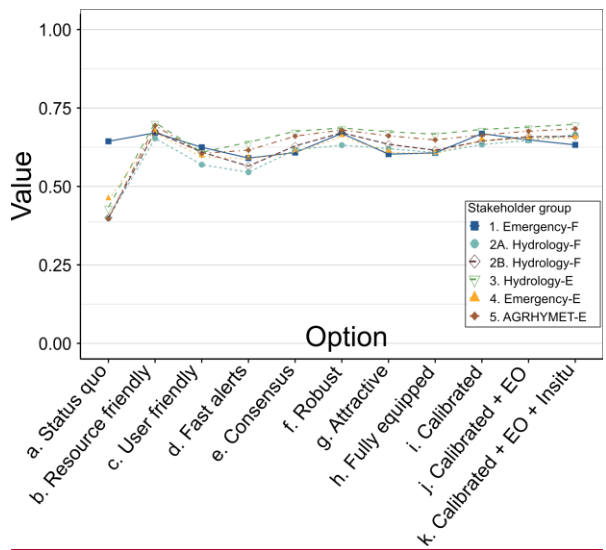



**Figure 5. Total aggregated value (y-axis) of 11 ~~FANFAR system~~FEWS configurations (x-axis) for six stakeholder groups (symbols), without uncertainty. Higher values indicate that they better achieved~~s~~ the objectives, given expert predictions and stakeholders' preferences.**

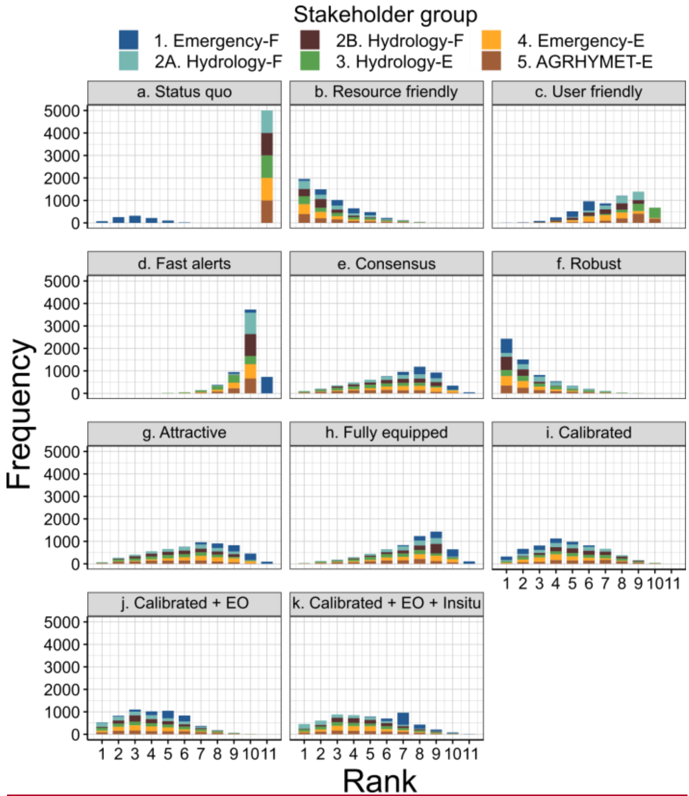



**Figure 6. Ranks of 11 ~~FANFAR system~~FEWS configurations including uncertainty of expert predictions. Frequency (y-axis): how often each ~~configuration~~ FEWS (blocks, *a. –~~Fast-dev~~Status quo*, *b. Resource friendly–~~Res-user~~*, etc.) achieved~~s~~ rank (1: best rank, 11: worst; x-axis) in each model run, for each stakeholder group (stacked bars). 1'000 Monte Carlo simulation runs drawing from uncertainty distributions of attribute predictions.**

Including the uncertainty of ~~the~~ expert predictions in ~~the~~ MCDA with Monte Carlo simulation clarified results. The ~~FANFAR system~~FEWS ~~configurations~~ *b. Resource friendly–~~Res-user~~* and *f. –Robust* performed well, ~~and~~ achieving~~ed the~~ highest ranks for all stakeholder groups in 1'000 simulation runs (Fig. 6~~Figure 6~~; details Table SI-34). The ~~configurations~~ FEWS *i. –Calibrated*, and *j. –Calibrated + -EO*, achieved good to medium ranks for most groups in most ~~simulations~~runs. Poor performance was achieved by ~~configurations~~ *a. Status quo–~~Fast-dev~~* (except group *1. Emergency-F*~~G1A~~), and *d. –Fast alerts*, which hit the last ranks in most simulation runs. The remaining ~~configurations~~ FEWS performed somewhere in between.

### ~~2.9~~4.7 Sensitivity analyses of stakeholder preferences

~~The performance of FANFAR system~~FEWS performance ~~configurations~~ was not sensitive to most model ~~parameter~~ changes (Table 2~~Table 3~~). The least changes in rankings occurred between the standard MCDA (S0) and sensitivity analyses ~~testing the~~ of extremes ~~of the~~ weight ranges elicited from stakeholders ~~in the workshop~~ (S22–S232; Table 2~~Table 3~~): Kendall's τ rank correlations were high, ranging from 0.86–~~to~~ 1 (1 = identical ranking of all ~~configurations~~FEWS). Doubling the weight of *23.*

*L*~~*l*~~*anguages* (S31) hardly impacted ~~the~~ rankings ~~for~~ of any stakeholder group. Greater changes occurred using other models~~aggregation models~~. The difference between the ~~our~~ standard MCDA (S0) and changed~~ing~~ aggregation models increased, the more the aggregation parameter $\gamma$ increased from 0 (geometric mean; S12), over mixture models (S13, S14), to 1 ~~the additive model with $\gamma = 1$~~ (additive model; S11). Rank correlations were still relatively high between the additive model ~~model~~ and ~~our~~ S0 ~~standard model~~ (0.53–0.86).~~, and~~ I~~i~~mportantly, ~~the~~ rankings of the best-performing ~~configurations~~FEWS, *b.* ~~–~~*Resource friendly*~~-user~~ and *f.* ~~–~~*Robust* did not change (~~details~~ sect. SI-2.8). For other configurations, including *i.* ~~–~~*Calibrated*, ~~there~~ some differences were ~~some~~ greater ~~differences~~, depending on ~~the~~ groups. The greatest changes occurred for ~~an~~ alternative weights ~~set~~ (S21) ~~elicited~~ in ~~the~~ group *1. Emergency-F*~~of French speaking emergency managers (G1A)~~. Interestingly, this ~~set~~ moved the rankings and values of ~~system~~ FEWS ~~configurations~~ to those of all other groups, this group ~~G1A~~ no longer being an outlier, and e.g., ~~configuration~~ *a*~~–Fast-dev~~*. Status quo* clearly performing worst ~~for all groups,~~ also for ~~-~~*1. Emergency-F*~~G1A~~ (~~Figure~~Fig. SI-40). Cost-benefit visualizations confirmed that ~~configurations~~ *b. Resource friendly*~~–Res-user~~, *f.* ~~–~~*Robust*, and *i.* ~~–~~*Calibrated* are suitable consensus ~~configurations~~ FEWS (see ~~provided in~~ sect. SI-2.9 for reasons of space).

**Table 2~~3~~. Results ~~of local~~ sensitivity analyses. Setting S0: default with elicited preferences of stakeholder groups and weighted power mean model, Eq. (3). Setting S11–S14: effect of other aggregation models (varying γ). S21–S22: uncertainty of Swing weights. S231–S232: uncertainty of Simos' card ~~method~~ weights. S31: increase (possibly underestimated) weight. S11–S31: all other parameters as S0. Columns group ~~G~~1~~A – to G~~5~~A~~: Kendall's τ rank correlation coefficient between ranks of ~~configurations~~ FEWS in main MCDA (setting S0) and ranks resulting from MCDA using other settings (S11–S31) for stakeholder groups (e.g., group~~-~~ *1. Emergency-F*~~G1A~~). Column mean: correlation between S0 and average rank over all groups for which analysis was done. Note: S21 was only done for group *1. Emergency-F*~~G1A–E~~ (i.e., mean = group correlation). Kendall's τ 1: identical ranks; 0: no correlation; –1: inverse relationship; –: not applicable. Kendall's τ from 0.81–1.00: underlined, indicating very good agreement between changed setting and S0; τ from 0.61–8.80: dotted underlined.**

| | | Stakeholder group | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Setting** | **Parameter change d~~D~~efinition** | **~~G~~1~~A~~** | **~~G~~2A** | **2~~G2~~B** | **~~G~~3~~A~~** | **~~G~~4~~A~~** | **~~G~~5~~A~~** | **Mean** |
| S0 | Default. MCDA for all six stakeholder groups; $\gamma = 0.2$; see Methods, eq. (3) | | | | | | | |
| S11 | Additive model all groups; $\gamma = 1$ | 0.86 | 0.64 | 0.60 | 0.64 | 0.53 | 0.75 | **0.67** |
| S12 | Weighted geometric mean all groups; $\gamma \to 0$ | 0.96 | 0.78 | 0.93 | 1.00 | 0.82 | 0.93 | **0.90** |
| S13 | Mixture model; $\gamma = 0.5$ | 0.93 | 0.78 | 0.67 | 0.75 | 0.75 | 0.75 | **0.77** |
| S14 | Weighted power mean; $\gamma = 0.8$ | 0.89 | 0.67 | 0.64 | 0.64 | 0.53 | 0.75 | **0.69** |
| S21 | Alternative weight set for group *1. Emergency-F*~~G1A~~ | 0.31 | – | – | – | – | – | **0.31** |
| S22_11_min | Weight ranges with $\Delta > 0.02$ from average weight for *11. Accuracy*~~–accur_info~~; minimum weight | 0.96 | – | 0.96 | – | – | – | **0.96** |
| S22_11_max | ~~*11_accur_info*~~*11. Accuracy*; maximum weight | – | – | 1.00 | – | – | – | **1.00** |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| S22_12_min | Weight ranges with $\Delta > 0.02$ from average weight for *1*2. *Clarity*~~2_clear_info~~; minimum weight | – | – | 1.00 | – | – | – | **1.00** |
| S22_12_max | *12*. *Clarity*~~_clear_info~~; maximum weight | 0.86 | – | 0.86 | – | – | – | **0.86** |
| S231 | Alternative weights ~~resulting~~ from ranges ~~assigned to~~of Z min | – | – | – | 0.89 | 0.89 | 0.89 | **0.89** |
| S232 | Alternative weights ~~resulting~~ from ranges ~~assigned to~~of Z max | – | – | – | 0.96 | – | – | **0.96** |
| S31 | Double weight of *23. Languages*~~"Several languages~~ | 0.96 | 1.00 | 0.93 | 0.93 | 0.89 | 1.00 | **0.95** |

### ~~2.10~~4.8 Stakeholders' perceived satisfaction with ~~the~~ current ~~FANFAR system~~FEWS

Participant numbers in the online workshop varied from 10–19 due to connection problems, which are frequent in West Africa, and related dropouts. The survey was filled out by 12 participants (12/19 = 63%), resulting for 10 objectives in 10 x 12 = 120 responses to each question. Most respondents perceived ~~the~~ current performance as sufficient for all objectives, based on the direct question about future use of the FANFAR ~~current FANFAR system~~FEWS (b), and the inferred difference (c ~~–~~minus a) between how much the ~~current~~ FEWS~~system~~ fulfills the respective objective (a) and the minimum acceptable level (c). Across all objectives, 79 responses were positive, 16 negative, and 25 did not answer question b. For the most important objective, *11*. ~~_a~~*Accur*acy~~_info~~, all respondents would use the current ~~system~~ FEWS in future (Fig. 7~~Figure 7~~). However, four (of 12) respondents indicated that the ~~system~~ FEWS does not currently meet their minimum acceptable performance requirements. This result is representative of ~~the~~ results for all objectives (details ~~see~~ sect. SI-2.10).

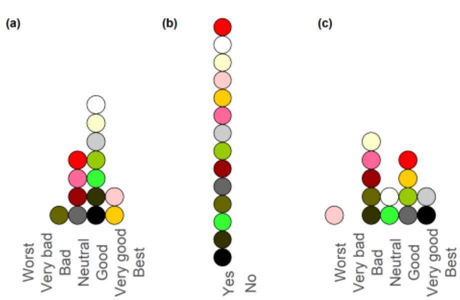



**Figure 7. Stakeholder perceived satisfaction with FANFAR FEWS ~~the~~ performance ~~of the FANFAR system during~~ in the 2020 rainy season for objective *~~11_accur_info~~11. Accuracy*. Questions: (a) How much does the ~~FANFAR system~~FEWS currently fulfil this objective? (b) Would you use the ~~system~~ FEWS in ~~the~~ future if it remains as is? (c) What is the minimum acceptable to you? Colored dots represent unique respondents (N = 12; 63% of 19 participants)~~)~~.**

### 4.9 Results conceptual framework for transdisciplinary process

The reviewed literature from sustainability science and transdisciplinary research revealed that many characteristics are shared between different authors (sect. 2.1). We included these in our proposed own framework (Table 3). However, some elements received less attention in this social science oriented literature, which are highly relevant to the earth system sciences: explicit consideration of uncertainty and the interdisciplinary effort needed for tackling technically complex problems (Lemos and Morehouse, 2005;Mauser et al., 2013). ~~The~~ Our framework follows a stepwise timeline, as proposed by many (Jahn et al., 2012;Lang et al., 2012;Mauser et al., 2013). We~~,~~ used~~ing~~ the terminology by Mauser et al. (2013): (i) co-design; (ii) co-production, and (iii) co-dissemination of knowledge (to which we added evaluation), involving academia and stakeholders throughout. For each phase, we included ~~specific~~ elements appropriate for ~~our context~~hydrology research for guiding ~~(Table 1).~~ and evaluating transdisciplinary processes.

**Table 3~~1~~. Conceptual framework for transdisciplinary research based on literature: (1) co-design, (2) co-production, and (3) co-dissemination of knowledge (terminology from Mauser et al., 2013), used to assess the MCDA process in the FANFAR project.**

| ID | Step | Explanation | Literature examples |
|---|---|---|---|
| **1** | **Co-design** | **Joint problem framing** | |
| 1a | Build collaborative research team | Include structures enabling participation from the start, e.g., use stakeholder mapping; aim at legitimacy of team; include bridging organizations or knowledge brokers to increase trust | (Lang et al., 2012;Wuelser et al., 2021;Norstrom et al., 2020) |
| 1b | Define research questions, methodological framework | Aim for balanced problem ownership from science and practice; define meaningful, shared goals, and measures of success | (Lang et al., 2012;Mauser et al., 2013;Jahn et al., 2012;Wuelser et al., 2021;Lemos and Morehouse, 2005;Norstrom et al., 2020) |
| 1c | Define boundary object | Translate problem into boundary object that allows re-integrating insights into societal implementation and scientific body of knowledge; “transformation knowledge” on how to make change e.g., with measures and tools | (Lang et al., 2012;Jahn et al., 2012;Schneider et al., 2019) |
| **2.** | **Co-production** | **Conducting integrated research to produce new knowledge; continuous exchange among scientists from different disciplines, and with stakeholders** | |
| 2a | Apply integrative (scientific) methods | Facilitate differentiation of different bodies of knowledge by using appropriate systematic procedures that ensure methodological consistency of research process | (Mauser et al., 2013;Wuelser et al., 2021;Lang et al., 2012;Jahn et al., 2012) |
| 2b | Interdisciplinary collaboration | Integrate knowledge of scientists from different disciplines; avoid conflicting methodological standards | (Mauser et al., 2013;Lemos and Morehouse, 2005;Jahn et al., 2012;Norstrom et al., 2020) |

| | | | |
|---|---|---|---|
| 2c | Explicitly consider uncertainty | Especially relevant in natural science problems addressing long time horizons (e.g., climate change) | (Mauser et al., 2013) |
| 2d | Integrate practice stakeholders in iterative process | Ensure appropriate roles, range of perspectives and skills, and context-based research; avoid discontinuous participation and vagueness of results that conceal potential conflicts | (Caniglia et al., 2021;Lang et al., 2012;Lemos and Morehouse, 2005;Norstrom et al., 2020) |
| 2e | Pluralistic principle/ social learning | Create shared understanding across multiple axes (e.g., disciplines, sectors, countries, gender); recognize values of people; foster training and capacity building | (Norstrom et al., 2020;Schneider et al., 2019;Caniglia et al., 2021;Wuelser et al., 2021) |
| **3** | **Co-dissemination and evaluation** | **Integrate and disseminate knowledge among research and societal groups in appropriate, relevant way; transparent discussion, critical reflection, and consequential actions** | |
| 3a | Two-dimensional integration | Review, discuss, and revise outcomes from societal and scientific perspective, e.g., prescriptive knowledge (recommendations about more desirable options) | (Lang et al., 2012;Caniglia et al., 2021;Mauser et al., 2013) |
| 3b | Generate targeted products | Translate results for scientific progress (e.g., generalizability), and real-world problem solving (e.g., relevance, scaling up results, alliances, actions in specific contexts, products such as maps, manuals, information for policy makers); knowledge transfer by scientists and societal actors | (Lang et al., 2012;Jahn et al., 2012;Wuelser et al., 2021;Caniglia et al., 2021;Lemos and Morehouse, 2005;Mauser et al., 2013) |
| 3c | Evaluate societal and scientific impact | Reference back to success factors (step 1b); impact can be defined in many ways, e.g., research quality, media attention, download rates, communities of practice, social networks, capacity building, education, concrete products, changing people's lives; longer-term impacts are often not measurable | (Lang et al., 2012;Jahn et al., 2012;Norstrom et al., 2020;Schneider et al., 2019;Lemos and Morehouse, 2005) |

# ~~3~~5 Discussion

The discussion follows the research questions (sect. 2.3~~1.3~~). RQA was confirmed: we found ~~a~~ robust ~~FANFAR system~~FEWS configurations despite large uncertainty and different stakeholder preferences (sect. 5.1~~4.1~~). We discuss our experience with MCDA ~~related to~~regarding uncertainty and eliciting stakeholder preferences (~~in~~ sect. 5.1.2~~4.1.2~~). To answer RQB, we presented main characteristics of transdisciplinary and sustainability science frameworks (sect. 4.9), proposing a framework for guiding and evaluating transdisciplinary processes (Table 3). Using our framework, ~~RQB was partly confirmed: early problem structuring focusing on objectives helped FANFAR system development (sect. 4.2). However, system developers needed MCDA weight elicitation asking for trade-offs to set priorities in the middle of the project, and full MCDA results to conclude that priorities were robust.~~ ~~i~~In RQC we analyzed the MCDA process for guiding large transdisciplinary projects (sect. 5.2~~4.3~~).

, following the proposed framework (sect. 2.1.1; Table 1). For step 1, co-design and joint problem framing, the FANFAR project met various requirements were met by the FANFAR project, but which could not be attributed to MCDA (sect. 5.2.14.3.1). However, if MCDA is broadly understood as a participative process that includes problem structuring, it can be very suitable for identifying stakeholders and guiding them to focus on objectives for achieving a joint boundary object, the FANFAR systemFEWS. For step 2, co-production of new knowledge, MCDA is appropriate (sect. 5.2.24.3.2). Core strengths of MCDA are allow that it can integratinge scientific knowledge from different disciplines by providingin a consistent framework and can handlinge uncertainty. MCDA invites stakeholders to clearly formulate their preferences , and identifies consensus FEWS configurations. Step 3, co-disseminating knowledge and evaluation, can only partly be achieved by MCDA (sect. 5.2.34.3.3). MCDA produces concrete, prescriptive knowledge: a suitable FEWS configuration of the FANFAR system. However, MCDA is not appropriate for other aspects such as producing and implementing real-world solutions, or impact evaluation. Main iInsights and recommendations are summarized in Table 4Table 4. and in the Conclusions (RQD, sect. 5).

## 3.15.1 Finding robust FANFAR FANFAR systemFEWS configurations (RQA)

### 3.1.15.1.1 Main MCDA results

As the most important practical result to RQA, we could identifyidentified three FANFAR systemFEWS configurations that hadwith a good overall performance (Fig. 5Figure 5). This would be difficult to achieve without MCDA, given the uncertainty of expert estimates and the model (Fig. 6Figure 6). Moreover, trade-offs between objectives had to be made (Fig. 3Figure 3), and stakeholders had different preferences concerning the importance of objectives (Fig. 4Figure 4). One well-performing configurationFEWS, *b. Resource friendly–Res-user*, was created by stakeholders in the first FANFAR workshop. They chose system FEWS components requiring the least resources for West Africa such as skilled personnel, good internet connection, or stable power supply (Table 1Table 2). Similarly, stakeholders created configuration *f. –Robust* was created by stakeholders to reliably work under difficult West African conditions related to collecting in situ data and distributing information through via a wide range ofvarious channels. The third configuration FEWS *i. –Calibrated* was created by FANFAR consortium members focusing on using refined HYPE models (e.g., adjusted delineation and parameter calibration; Andersson et al., 2020b), but not includingexcluding earth observation andor in situ data (included in FEWS configuration *j* and *k;* Table 1Table 2). All three best configurations FEWS achieved 63–70 % of all objectives in all stakeholder groups. We consider this a very good value, given the existing trade-offs. These configurations FEWS also emerged aswere robust consensus from (i) including (i) when including the uncertainty of expert predictions with Monte Carlo simulation (Fig. 6Figure 6); (ii) in sensitivity analyses of the aggregation model assumptions and changing stakeholders' weight preferences (Table 2Table 3); and (iii) in dominance checks in cost-benefit visualizations (sect. SI-2.9). Interestingly, none of these three configurations FEWS did not incorporate more advanced system features: a . Hence, a flood forecast and alert systemFEWS that meets stakeholder the preferences of the West African stakeholders primarily needs to work accurately and reliably under difficult West African conditions in West Africa.

### 3.1.25.1.2 Dealing with uncertainty of predictions, preferences, and model assumptions

Attributes operationalize objectives objectives (Eisenführ et al., 2010). Seemingly trivial, this is often challenging. We illustrated this for 11_kge, constructed thefrom the KGE index for 1, 3, and 10 day forecasts to measure achieving objective *1.1. High aAccuracy of information* (sect. 3.1.52.2.7). The uncertainty of expert predictions was relatively large for for some attributes (e.g., *11_kge11. Accuracy*, *22. –Ttimeliness*, or *42–experts. Skilled labor*), but small to inexistent for others (e.g., *12. Clarity–info*, and *23. –Languageslangue* (; FigureFig. SI-30). The resulting overall uncertainty affected the results less than might be expected (FigureFig. SI-35).

The weights indicated that most groups preferred a system FEWS producing accurate, clear, and reliable information, reaching recipients well before floods (*11. Accuracy, 12. Clarity, 21. –accur_infoAccess; 12_clear_info; 21_reliable_info*; *22. –timely infoTimeliness*; Fig. 4Figure 4), and West African countries need the capability to handle this information (*42. –human_resourSkilled labor*). We captured differences within groups with uncertainty ranges or separate preference sets (e.g., subgroups *G1A, G2B2A, 2B*; ; sect. SI-1.2.3; sect. SI-2.6). The French speaking emergency managers (*1. Emergency-F*(G1A) had different preferences compared to all others. All groups regarded several languages as unimportant in weight elicitation, despite discussing in the 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 latterlanguage diversity to achieve accuracy. They were also willing to trade-off higher operation and maintenance costs (except *1. Emergency-F*group G1A) and development time in return for receiving a functioning, precise FEWSsystem.

Including the uncertainty of expert estimates and stakeholder preferences in MCDA can blur results. For FANFAR, Multi-Attribute Value Theory and local sensitivity analyses (e.g., as Zheng et al., 2016) enabled identifying configurations suiting all stakeholder groups. iIncluding the uncertainty of predictions helped to *better* distinguish between FEWS performances of configurations (Fig. 6Figure 6), compared to the standard analysis without uncertainty (Fig. 5Figure 5). Configurations FEWS configurations *b. Resource friendlyb_Res-user* and *f. –Robust* consistently achieved the first ranks in 1’000 simulation runs, and e.g., *i. –Calibrated* good to medium ranks. However, some configurations FEWS such as *k. Calibrated k–Cal+ -EO + In-situ*, ranked last in numerous runs (Fig. 6Figure 6), despite achieving good values when uncertainty was disregarded (0.63–0.70; Table SI-33). Ranking last in most runs, *a. –FastStatus quo-* and *d. –Fast alerts–dev* would be an imprudent choice.

Llocal sensitivity analyses (e.g., as Zheng et al., 2016) enabled identifying FEWS suiting all stakeholder groups. Local sensitivity analyses confirmed that configurations *b. Resource friendlyb–Res-user*, *f. Robustf–Robust*, and *i. Calibratedi–Calibr* are robust choices. Changing stakeholder preferences hardly changed MCDA results compared to our standard model (S0; Table 2Table 3). Doubling the weight of *23. –Llanguages* (S31) did not affect results in any group, thus avoiding costly translations as priority. Operation and maintenance costs would have been another candidate for doubling the weight, but was covered by the high weight of group *1. Emergency-F*G1A_EM_F. In this group, sensitivity analyses on weight ranges given by group participants with a different opinion (S21; Table 2Table 3) changed the results so that they aligned with results of the other stakeholder groups. This increases our confidence that the three proposed configurations FEWS are a good consensus. Moreover, the additive MCDA aggregation model (Eq. (1); sect. 3.1.72.2.9) impacted the FEWS rankings of the FANFAR system configurations (Table 2Table 3). As standard model, we assumed non-additive aggregation (Eq. 3), which is relatively close to a weighted geometric mean model, based on feedback in the weight elicitation sessions. After discussing examples, all groups stated that poor performance on an important objective should not be compensated by good performance on others, a main implication of additive aggregation. This confirms that the additive model can unintentionally violate stakeholder's preferences (e.g., Haag et al., 2019a;Reichert et al., 2019;Zheng et al., 2016). Thus, additive aggregation may not be the best model, despite its popularity in MCDA applications. For FANFAR, sensitivity analyses sufficed to conclude that additive aggregation has an effect, but does not alter rankings of the best configurationsFEWS. We can thus safely conclude that the three proposed FANFAR systemFEWS configurations are suitable. We emphasize that the FANFAR systemFEWS was continuously improved throughout the project, also after eliciting stakeholder preferences in the second workshop.

### 3.2 Early problem structuring focused FANFAR system development (RQB)

**3.3 Early problem structuring steps helped focusing technical system development before MCDA results were available (RQB) This is important for agile development processes, since complex MCDA often takes considerable time to produce solid results. We emphasized objectives, rather than starting with FANFAR system configurations, motivated by "Value focused thinking" (Keeney, 1996). Asking: "What is important when designing a flood forecast and alert system for West Africa?" in the first workshop allowed system developers to focus on priority objectives. Elicitation of stakeholder preferences in the second workshop further guided system design, e.g., that not all system components needed translation to several languages (sect. 4.1.2). Moreover, designing a flood forecast and alert system is complex (Andersson et al., 2020b;Arheimer et al., 2011;Emerton et al., 2016), with a risk of getting lost in stakeholder discussions about system elements. To avoid this, we included all stakeholder suggestions when creating system configurations. Hydrologists and ICT specialists then aimed at finding best solutions, given the stakeholders' priorities and experts' predictions about system performance. To conclude, early problem structuring was useful, but later weight elicitation was needed to prioritize subsequent FANFAR system development. The MCDA results confirmed the robustness of mid-term priorities, which can be different in each case. We thus recommend doing the entire MCDA to test the sensitivity of results (sect. 4.1.2). The stepwise iterative approach employed here (beginning with key objectives, then focusing on system configurations), was very productive, and would likely be useful also in other projects.**

### 3.45.2 Suitability of the MCDA process for guiding large transdisciplinary projects (RQC)

We discuss research question RQC following the framework for transdisciplinary research (Table 1). We critically evaluate a participatory MCDA process for guiding a large transdisciplinary research project following our proposed framework (Table 3). We focus on on aspects that we consider important aspects when carrying outof MCDA in a hydrology context,. summarizing mMain points and recommendationsin are summarized in Table 4Table 4.

#### 3.4.15.2.1 Evaluating the co-design step "joint problem framing"

The MCDA process does not fully meet all requirements of this step. ***In the FANFAR project, bBuilding the collaborative research (or project) team*** - with consortium partners from Europe and West Africa cannot be attributed to MCDA, although it was achieved by the FANFAR project (step 1a,- Table 3Table 1), but not by MCDA.. Two Two key West African stakeholders were consortium partners from the start: : AGRHYMET (mandated by 13 West African states and ECOWAS to provide e.g., operational flood warnings), and NIHSA (Nigerian Hydrological Services Agency). This follows a decade of collaboration between SMHI and AGRHYMET. Building alliances with regional partners is a transdisciplinary approach identified across projects, and may lead to follow-up ppartnerships in follow-up work (Wuelser et al., 2021). ***Trust building*** is crucial (e.g., Lemos and Morehouse, 2005), and AGRHYMET is clearly a ***bridging organization or knowledge broker*** between research and implementation (Norstrom et al., 2020;Wuelser et al., 2021;Lemos and Morehouse, 2005) between research and implementation. The FANFAR project was co-led by West African partners and engaged stakeholders in workshops, meeting the principle of creating ***knowledge tailored to specific contexts*** (Caniglia et al., 2021;Norstrom et al., 2020). However,, but this cannot be attributedis not due to MCDA, nor. Similarly, MCDA was not responsible for ***defining the research questions*** (step 1b, Table 1), nor or the ***boundary object*** (step 1c, Table 3). The FANFAR had a concrete boundary object was to produce an operational FEWS, which allowedalloweding stakeholders to commit (Jahn et al., 2012): producing an operational flood forecast and alert system. SIt was in the interest of scientists and stakeholders both aimed to achieve this goal, thus which helpeding to overcominge the problem of unbalanced ownership (Lang et al., 2012). The FANFAR consortium agreed to use MCDA as ***integrative methodological framework*** (1b) for to achieveing this goal objective. Literature emphasizes the importance of using methodologically consistent procedures forand integrateing across different scientific disciplines (Lemos and Morehouse, 2005;Mauser et al., 2013;Lang et al., 2012). We regard MCDA is oneas such apossible useful, stringent, and integrative methodology. Indeed, stakeholder related multi-criteria assessment (i.e., MCDA) was mentioned as one possible methodology to produce ***transferable knowledge*** (Wuelser et al., 2021).

Narrowing the perspective to the concrete project in with West African with local stakeholders, an MCDA process that emphasizinges early problem structuring is helpful (Marttunen et al., 2017;Rosenhead and Mingers, 2001) can be helpful. ***Taking practitioners on board*** from the beginning start and avoiding insufficient legitimacy or underrepresentation of actors is crucial (Wuelser et al., 2021), and insufficient legitimacy or underrepresentation of actors is a challenge (Lang et al., 2012;Wuelser et al., 2021). To identify those that should be involved, sStakeholder mapping or social network analysis are suitable to identify those to involve (Norstrom et al., 2020;Lang et al., 2012). As first step of MCDA, we carried out a ***stakeholder analysis*** (step 1a,- Table 3Table 1), which is rarely . However, in typical MCDA projects, stakeholder analysis was done in only 9% of 333 reviewed papers (9% of 333 reviewed MCDA papers; Marttunen et al., 2017). We used relatively simple questionnaires (sect. 3.1.22.2.4) to discover who has influence or power or, and who is affected by a good or malfunctioning flood forecast and alert systemFEWS (Grimble and Wellard, 1997;Lienert et al., 2013;Reed et al., 2009). We identified 68 distinct stakeholder types (sect. 4.13.1). In the workshops, we included hydrologistsy representatives from 17 countries, and key supranational organizations such as AGRHYMET who produce flood information (Table SI-4; details see Silva Pinto and Lienert, 2018). Main receivers of FEWSflood information also participated: emergency managers from every country. Thanks to their experience, we integrated the alert dissemination chain in FANFAR and elaborated the alert dissemination chain and elements of effective flood early warning systemsFEWS (Kuller et al., 2021). We also identified missing parties, e.g., agriculture, industry (dam managers, electricity utilities), or humanitarian aid organizations., Ssome of which provided informal feedback on the FANFAR systemFEWS through social media. We did not invite them because more than 50 participants in workshops is ineffective. Indeed, ***pluralistic co-production*** while ***keeping processes manageable*** remains a challenge (Norstrom et al., 2020;Lang et al., 2012).

***Problem structuring*** is decisive because MCDA results critically depend on objectives and options (i.e., FEWS configurations Marttunen et al., 2017;Rosenhead and Mingers, 2001). These MCDA steps (Figure 1) were carried out in the first stakeholder workshop (Fig. 1; sects. 3.1.32.2.5, and 3.1.42.2.6). At a specific level, the interactionsThey helped define meaningful, ***shared goals, and one a success measure of success*** (e.g., Norstrom et al., 2020; step 1b, Table 3): to find a FANFAR systemFEWS configuration that meets achieves the objectives that should be achieved by this system. Following V"Value Ffocused Tthinking" (Keeney, 1996), we first generated objectives in small groups using different methods (sect. 3.1.32.2.5). This to ensured a a broad diversity of objectives and helped avoid the common "group think bias" (Janis, 1972). We are confident that we captured the most important 10 objectives that cover fundamental aims of West African stakeholders (Fig. 2Figure 2). Moreover, many environmental applications of MCDA use too many objectives (Marttunen et al., 2018). This is ineffective and burdens MCDA weight elicitation. We excluded some objectives in plenary discussions.

We could not assume that all participants had sufficient technical knowledge to ***generate create system configurationsFEWS configurations***, but aimed to avoid "myopic problem representation", a bias to stick to what one knows (Montibeller and von Winterfeldt, 2015). We find tThe Strategy Generation Table is especially suitable (Gregory et al., 2012b;Howard, 1988) especially suitable. It allows pre-structuringes the process while while allowing forstimulating creative stakeholder inputs. The ***context-based principle*** of co-production includes asking for constraining factors (Norstrom et al., 2020): when . When creating FANFAR systemFEWS configurations, it the necessity of became evident thatconsidering the West African situation became evident, including e.g., frequent power cuts and slow internet in West Africa need consideration. Moreover, we realized that stakeholders had not created all potentially interesting FEWS configurations. An advantage of Multi-Attribute Value Theory (Eisenführ et al., 2010) is that allows later including system options can be included later configurations (Reichert et al., 2015;Eisenführ et al., 2010). The The FANFAR consortium created additional configurations FEWS to covering technical aspects, e.g., ensemble meteorological forecasts, redelineation and calibration of hydrological models, and assimilation of EO and in situ water levels and in situ gauge observations from rivers (configurations FEWS *h* to *k*, Table 1Table 2; Table SI-6). The During post-processing, we also created the status quo systemFEWS at project start, *a. Status quo*,a_Fast-dev, was also created during post-processing to serve as benchmark. Indeed, it performed poorly for most groups (Fig. 5Figure 5). As summary, the three MCDA steps of stakeholder analysis, creating objectives, and system configurationsFEWS took up a large parts of the first West African workshop. They were very helpful for stakeholders to exchange ideas, express their needs and wishes, and ***develop a common understanding***, contributing to the co-design step 1 (Table 4Table 4).

**Table 4. Summary of MCDA process using conceptual framework for transdisciplinary research (Table 3~~Table 1~~): (1) co-design, (2) co-production, and (3) co-dissemination of knowledge. Symbols: +++ strength of MCDA; ++ well possible with MCDA; + possible contribution by MCDA; 0 not achievable by MCDA; * remark. PSM: Problem Structuring Methods; VFT: Value Focused Thinking.**

| ID | Step | MCDA | Remarks and recommendations |
|---|---|---|---|
| **1** | **Co-design** | **Joint problem framing** | |
| 1a | Collaborative research team | 0 | • Include local partners in consortium (knowledge brokers, bridging organizations) |
| | | 0 | • Build alliances with regional partners, also for follow-up projects (trust building) |
| | | +++ | • MCDA PSM: stakeholder analysis with simple questionnaires (sect. 3.1.2~~2.2.4~~) |
| 1b | Research questions, methodological framework | ++ | • MCDA can help jointly defining research questions if PSM is used |
| | | 0 | • MCDA is less suitable to define project success criteria (but PSM could be used) |
| | | +++ | • MCDA provides an integrative methodological framework (sect. 3.1~~2.2~~) |
| | | ++ | • MCDA PSM: use VFT for defining shared objectives at lower level (sect. 3.1.3~~2.2.5~~) |
| | | ++ | • MCDA PSM: use creativity techniques to find diverse, locally adapted solutions (e.g., ~~strategy~~ Strategy ~~building~~ Generation T~~t~~able; sect. 3.1.4~~2.2.6~~); increases common understanding |
| 1c | Boundary object | + | • MCDA PSM could potentially be used for creating boundary object |
| **2** | **Co-production** | **Conducting integrated research to produce new knowledge; continuous exchange** | |
| 2a | Integrative methods | +++ | • MCDA is a methodologically consistent integrative procedure, but there are others |
| 2b | Interdisciplinary collaboration | +++ | • MCDA can integrate qualitative and quantitative scientific evidence from different disciplines using predictions (sect. 3.1.5~~2.2.7~~) and value functions (sect. 3.1.6~~2.2.8~~) |
| | | | * Not emphasized in reviewed transdisciplinary literature: merits future research |
| | Uncertainty | +++ | • MCDA can explicitly consider various types of uncertainty (sect. 5.1.2~~4.1.2~~) |
| | | | * High relevance for projects in the earth systems sciences; merits future research |
| 2d | Integrate practice stakeholders | ++ | • MCDA can integrate diverse practice stakeholders throughout project |
| | | 0 | • Iterative process to integrate practice stakeholders should be included in MCDA |
| | | 0 | • MCDA cannot handle discontinuous participation, or too many participants |
| | | +++ | • MCDA provides clear results, avoiding vagueness that conceals potential conflicts |
| 2e | Pluralistic principle/ social learning | +++ | • MCDA explicitly recognizes different stakeholder interests, which are integrated in model, fostering trust and avoiding conflict by finding consensus configurations |
| | | +++ | • MCDA weight elicitation: allow for uncertainty & different stakeholder preferences |
| | | +++ | • MCDA fosters learning about decision, one's own preferences, and those of others |
| | | | * To understand growing shared understanding in a group, future research is needed |

| | | | |
|---|---|---|---|
| | | 0 | • MCDA does not foster training and capacity building |
| **3** | **Co-dissemination** | | **Integrate and disseminate knowledge among research and societal groups, and evaluation** |
| 3a | Two-dimensional integration | ++ | • MCDA: some discussion and revision of results (to find consensus FEWS configurations) |
| | | +++ | |
| | | 0 | • MCDA provides prescriptive knowledge (e.g., suitable system FEWS configurations) |
| | | | • MCDA cannot review and analyze other aspects (e.g., governance mechanisms) |
| 3b | Targeted products | 0 | • MCDA cannot generate target products (e.g., publications, policy briefs, maps) |
| | | 0 | • MCDA cannot implement and scale up knowledge for real-world problem solving |
| | | + | • Scientific integration, generalization, and documentation is not specific to MCDA |
| 3c | Evaluate societal and scientific impact | + | • MCDA usually does not evaluate societal and scientific impact (but is possible) |
| | | 0 | • Mid-term impacts cannot be attributed to MCDA (e.g., uptake, societal effects) |
| | | 0 | • MCDA cannot capture longer-term impacts, which are anyway difficult to measure |

### 3.4.25.2.2 Evaluating the co-production step "integrated research to produce new knowledge"

Using cConsistent ***integrative methods*** and ***systematic procedures*** for integrating bodies of knowledge are crucial (step 2a, Table 3Table 1) and ***systematic procedures*** for different bodies of knowledge (Lang et al., 2012;Mauser et al., 2013) is crucial, but seems less visible in literature (Wuelser et al., 2021;Lang et al., 2012;Mauser et al., 2013). Recommendations include generating hazard maps, qualitative system analysis, scenario analysis, and or sensitivity or and multi-criteria assessments (i.e., MCDA). Applying procedures to iIdentifying stakeholders' positions and preferred configurations options is beneficial because itallows involvinges people in creating their future (Wuelser et al., 2021). In FANFAR, MCDA clearly helped structuring the co-design process and integrating different types of knowledge types: expert estimates of how well each FANFAR systemFEWS configuration achieves each objectiveperforms (sect. 4.43.4) and stakeholders' preferences (sect. 4.53.5). At Moreover, each workshop, West African stakeholders experimented with the FANFAR systemFEWS at each workshop, tested it in rainy seasons, and provided feedback (see Wuelser et al., 2021), which cannot be attributed to MCDA.

Transdisciplinary projects rely on ***interdisciplinary collaboration*** and integrating evidence from different disciplines (step 2b, Table 3; Jahn et al., 2012;Lemos and Morehouse, 2005;Mauser et al., 2013). Methods must allow for consistent integration of scientific evidence from different disciplines to avoid "comparing apples with pears" (Mauser et al., 2013). Integrating qualitative data for policy and decision making, and quantitative data required byfor models can be challenging (Lang et al., 2012). MCDA handles this by transforming attributes of different measurement units (including qualitative scales) to a common scale value between from 0 (objective not achieved) to 1 (objective fully achieved), using value functions (sect. 3.1.62.2.8). In FANFAR, experts from different disciplines provided these estimates: West African and European hydrologists, IT specialists, and decision analysts (sect. SI-2.4.1). MCDA integrates very specific data (predictions about FEWS performance of system configurations). Other evidence types of evidence also need integration in transdisciplinary projects, and other methods are available. This area merits future research, given the lack of emphasis in current literature.

"Questions of the ***uncertainty of the results***" (Mauser et al., 2013; p. 428) were emphasized by earth systems scientists for global sustainability, but scarcely addressed by others (step 2c,- Table 3Table 1). Our MCDAWe included the uncertainty of expert predictions by eliciting probability distributions for each attribute (sect. SI-2.4.1) and Monte Carlo simulation (sect. 3.1.82.2.10). Local sensitivity analyses addressed uncertainty of the model and of stakeholder preferences (sect. 5.1.2; discussed in Reichert et al., 2015). Handling uncertainty in a conceptually valid way is essential for transdisciplinary research in the earth systems science.

The importance of ***integrating practice stakeholders in iterative processes*** (step 2d,- Table 3Table 1) was underlined by many (e.g., Lemos and Morehouse, 2005;Norstrom et al., 2020). Our iterative workshop series to test and improve the FANFAR systemFEWS (sect. 4.1.2) cannot be attributed to MCDA. Practical MCDA projects often consist of three stakeholder interac-tionsworkshops: a first workshop for problem structuring, a second for preference elicitation, and a third for discussing results and revising options, i.e., system FEWS configurations ((Fig. 1)Figure 1). ***Discontinuous participation*** can be a challenge (Lang et al., 2012), also for FANFAR, and . The FANFAR faced 3.5-year project involved representatives of at least two institutions from 17 African countries, regional stakeholders, and European partners, reflected in changing numbers and composition of participants (sect. 3.1.12.2.3). As Lang et al. (2012), we also encountered the opposite: increasing requests over time and the challenge of ***keeping participant numbers manageable***. We integrated new participants, e.g., by presenting the FANFAR systemFEWS and MCDA objectives at each workshop. For MCDA, discontinuous participation was unproblematic, as new participants in the second workshop accepted the objectives (Fig. 2Figure 2) and system configurationsFEWS (Table 1Table 2). This indicates that oOur participant sample was presumably sufficiently large and diverse to cover the main aspects. Another challenge can be ***vague results*** when using methods such as sustainability visions, which may ***conceal potential conflicts*** (Lang et al., 2012). MCDA has the strength of providing clear results, even for uncertain data (sect. 5.1.24.1.2).

The ***pluralistic principle aims at creating social learning across multiple axes*** (step 2e,- Table 3Table 1). , and to build trust, e.g., by recognizing values of people, providing trainings, and capacity building. Sustained interaction with stakeholders, jointly searching for solutions, joint practices, and joint learning foster ***trust***, ***mutual understanding***, and ***shared perspectives*** (e.g., Lemos and Morehouse, 2005;Norstrom et al., 2020;Schneider et al., 2019). , and "relationships of trust among partici-pants make policymaking and implementation more likely to succeed" (Lemos and Morehouse, 2005; p. 61). ***Recognizing different expertise, perspectives, values, and interests*** does not require reaching consensus (e.g., Norstrom et al., 2020;Wuelser et al., 2021). does not require consensus to be reached across conflicting interestsHowever, , but implies collaborativelyly ***engaging with conflicts*** (Caniglia et al., 2021)is needed to rationalize contested situations (Schneider et al., 2019;Caniglia et al., 2021). A strength of MCDA is that opposing stakeholder interests are part of the methodology, hereby often avoiding conflict about the choice of solutions (Arvai et al., 2001;Gregory et al., 2012a;Gregory et al., 2012b;Marttunen and Hamalainen, 2008). In theDuring weight elicitation sessions, we encouraged stakeholders to discuss diverging preferences about the importance of objectives (sect. 3.1.62.2.8), and we recommend allowing for such uncertainty. It helps participants ***construct their own preferences*** (Lichtenstein and Slovic, 2006), enables ***learning*** and ***understanding alternative perspectives***, and can informs sensitivity analyses (sect. 4.73.7). In FANFAR, conflicting preferences did not change the ranking of system configurationsFEWS rankings, and we identified consensus systems FEWS (sect. 5.1.24.1.2). In other cases, sensitivity analyses based on diverging preferences can help construct better configurationsFEWS. As interesting observation, in the first FANFAR workshop, participants voted on objectives to include in the MCDA, which did not match an African perspective for all. Some participants criticized the democratic approach, remarking e.g., that "the chief should decide". "Assessing the [interactive] principle should also focus on capturing learning, how the perceptions of actors change throughout the process, and the degree to which a ***shared perspective emerges***" (Norstrom et al., 2020; p. 188). Such research is rarely done in MCDA, but was attempted in FANFAR and a Swiss project. We shortly elicited the importance of objectives at every workshop, hypothesizing that increased shared understanding is mirrored by converging stakeholder preferences (Kuller et al., in prep.). Results were ambiguous, but we found shared agreement of FANFAR stakeholders about the most important objectives. More research to better understand individual cognitive and group decision making processes is needed (Kuller et al., in prep.).

***Training and capacity building*** are components ofbelong to the ***pluralistic principle*** (step 2e, Table 3Table 1). To foster joint understanding, commitment, and trust, mMany 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 co-production discussed above (Caniglia et al., 2021), or with capacity building courses (Wuelser et al., 2021;Caniglia et al., 2021). The FANFAR project offered many training and capacity building opportunities, which cannot be attributed to MCDA.

### 3.4.35.2.3 Evaluating the co-dissemination and evaluation step "integrating and disseminating knowledge"

***Two-dimensional integration*** (step 3a, Table 3Table 1) implies that project outcomes are ***reviewed, discussed, and revised*** from a scientific and societal perspectives (Mauser et al., 2013;Lang et al., 2012). Discussing ***transformation knowledge*** includes measures, tools, or governance mechanisms to create change in the desirable direction (Schneider et al., 2019). It can include ***prescriptive knowledge***, recommending suitable options for realizing intentions (Caniglia et al., 2021). This is a strength of MCDA: in our case, we provided detailed information about good, robust FANFAR systemFEWS configurations (sect. 5.14.1). Moreover, MCDA results are discussed with stakeholders and if needed, new consensus optionsFEWS could be are constructed (Fig. 1)(Figure 1). We were could not able to carry out the fourth FANFAR workshop in West Africa due to the COVID-19 pandemic, but. We collected some online feedback about how well the current FANFAR system meets the 10 objectives in an online survey. The sStakeholders were quite satisfied with the FANFAR FEWS its performance during the 2020 rainy season (Fig. 7)(Figure 7), and are willing to use it in future. Although tWhilehis does not meeting requirements of extensive discussions, itit was the best available approach. We are currently carrying out a systematic daily reforecasting experiment covering 1991–2020 for five model configurations, and aim to link results to expert satisfaction. Understanding

***governance mechanisms*** is out of ~~the~~ scope of MCDA; in our case, ways to facilitate uptake of the ~~FANFAR system~~FEWS across entire West Africa.

***Target products*** (step 3b,~~-~~ Table 3~~Table 1~~) should address the original problem, be understandable, and accessible to users (Lemos and Morehouse, 2005;Schneider et al., 2019;Lang et al., 2012). Products include technical publications ~~in appropriate language~~, data visualizations, and open access online databases (Schneider et al., 2019). In FANFAR, ~~these~~ products cannot be attributed to MCDA. Main product is the ~~operational FANFAR flood forecast and alert system~~FEWS (including operational data collection, assimilation, hydrological modelling, interpretation, and distribution through web visualization and API), where MCDA only ~~helped in the~~supported the design. Additional products ~~include~~ are a multilingual ~~open access~~ knowledge base (https://fanfar.eu/support/), ~~the~~ open source code (https://github.com/hydrology-tep/fanfar-forecast), and video tutorials (www.youtube.com, search: HYPEweb FANFAR). Assuring ***consistent access, maintenance, updates, and improvements*** after project termination is ~~a~~ challenging~~e~~ (Lemos and Morehouse, 2005). AGRHYMET has the authority to drive the FEWS uptake ~~of FANFAR in West Africa~~ and already uses it, e.g., ~~beyond project activities~~ in their MSc curriculum~~, e.g.~~, or at PRESASS and PRESAGG forums (WMO, 2021), ~~to support~~supporting the ECOWAS flood management strategy~~, and in their MSc curriculum~~. Nevertheless, ~~long term~~ operationalization after EU financing is not ~~guaranteed~~secured.

***Products should contribute to scientific progress*** ~~(Jahn et al., 2012)~~, a major challenge being inadequate ***generalization of case study solutions*** (Lang et al., 2012;Jahn et al., 2012). Products are often not reported in scholarly literature (Wuelser et al., 2021), the knowledge thus not advancing scientific progress, and not being adopted in ~~societal praxis of~~ similar projects. We aimed to overcome this with this paper and other outputs (FANFAR, 2021). ~~For MCDA, w~~We document the MCDA process, ~~and~~ providing~~e~~ details in the Supplementary Information. We encourage hydrologists to use this material. We stress that it is not necessary to conduct a full MCDA in every case. The first p~~P~~roblem structuring steps ~~(sect. 4.3.1)~~ can create useful insights ~~(sect. 4.2)~~, and may be easier to apply (sect. 5.2.1).

The last step 3c (Table 3) ~~(Table 1)~~ is to ***evaluate societal and scientific impact***. ~~MCDA did not contribute to this, although p~~Project evaluation ~~would be~~is possible with MCDA, but MCDA was not used in FANFAR. Short-term impacts ~~can be visible in~~include ***increased citations or attention of nonacademic actors***, e.g., high ~~including~~ download rates, or media coverage (Norstrom et al., 2020;Schneider et al., 2019). ~~This occurred; a~~As example, the ~~third~~ FANFAR workshop in Nigeria featured on the national TV news. ***Building social capacities and establishing stakeholder networks or communities of practice*** can be very helpful (Lemos and Morehouse, 2005;Schneider et al., 2019). A FANFAR social media group among ~~the~~ West African stakeholders monitored the severe 2020 floods ~~across the region~~, which in many places were successfully forecasted by the FANFAR ~~FANFAR system~~FEWS. ~~Indicators of m~~Mid-term impact include~~s relate to~~ ~~the~~ ***uptake of*** ~~***concrete***~~ ***products*** and ~~to~~ societal effects ~~of the transdisciplinary process~~ such as ***strategy implementation, or amended legislation*** (Jahn et al., 2012;Norstrom et al., 2020). ***Long-term impacts*** are very difficult to measure as they are typically realized far beyond project termination (Norstrom et al., 2020;Schneider et al., 2019). Moreover, due to the complexity of ~~the~~ problems ~~addressed~~ in transdisciplinary projects, ~~it is difficult to establish~~ causal relationships are difficult to establish (Lang et al., 2012). ~~Whether~~

the FANFAR flood forecast and alert system will be taken up broadly by West African stakeholders is uncertain. To secure future sustainability of the FANFAR FEWS, a set of dialogues with potential financiers were held, and 12 proposals were submitted to date. Four were successful so far, providing funding for some parts of FANFAR (e.g. hydrometric stations by AfDB, additional training by Sida and EDF via ECOWAS). The sustainability strategy focuses on financing (of operations, maintenance, dissemination, technical development, etc.) and importantly on long-term collaboration, capacity development, transfer of responsibilities, and on anchoring FANFAR However,in the routines of West African institutions. it already had As one example of ***societal impact***. As one example, the NIHSA (Nigeria Hydrological Services Agency, NIHSA) reported that an early FEWS warning received from the FANFAR system in September 2020 saved approximately 2’500 lives. The warning helped evacuating five communities before the flood destroyed more than 200 houses.

## 46 Conclusions

The MCDA process enabled identifying finding three good FANFAR FANFAR systemFEWS configurations (Research Question RCA; sect. 4.1), which is important to West African stakeholders and peoplethose potentially affected by floods. These consensus configurations performed well for all stakeholders, despite their diverging priorities. All stakeholder groups stakeholder groups preferred a relatively simple system FEWS producing accurate, clear, and accessible flood risk information that reaches recipients well before floods. To achieve this, most groups would trade-off higher operation and maintenance costs, development time, and several languages. The West African stakeholders preferred a system that works reliably and accurately under difficult conditions in the region, but not necessarily including many advanced features. MCDA demonstrated indicated the robustness of thesethat the three configurations FEWS are robust. They achieved 63–70% of all 10 objectives despite given differentdiverging stakeholder perspectivespreferences, model uncertainty, and uncertain expert predictions. from different disciplines concerning FANFAR system performance (sect. 4.1.2). Including uncertainty and stakeholders in MCDA is neglected in flood risk research. The three FANFAR system configurations achieved 63–70 % of the entire set of objectives for all stakeholder groups. Interestingly, the uncertainty analysis allowed to better distinguish between system configurations than standard MCDA without uncertainty. We highly recommend both: MCDA including uncertainty allowed better distinguishing between FEWS, and participatory MCDA focusing on stakeholders’ objectives ( using a participatory MCDA process to find suitable options in hydrology researchValue Focused Thinking), . Emphasizing objectives follows Value Focused Thinking (VFT), which helped avoiding potential conflicts about system FEWS configurations. Hopefully, participatory MCDA increased trust and capacity building among West African stakeholders, thus enabling future uptake of the FEWS. Important for practice, VFT was also beneficial to guide early technical FANFAR system development, before MCDA results were available. However, eliciting the trade-offs that stakeholders were willing to make was necessary for later setting priorities, and the robustness of results could not have been evaluated without the full MCDA (RQB; sect. 4.2).

~~From the perspective of sustainability science and transdisciplinary research, ~~MCDA meets many, but not all requirements of sustainability science and transdisciplinary research~~(RCD; Table 4)~~. ~~The ~~Our proposed evaluation framework ~~based on literature ~~proved very useful for ~~our ~~critically analyzing~~sis of~~ MCDA~~ (Table 1)~~. We invite others to apply it for guiding their projects, and to ~~apply and ~~evaluate our framework~~it in their cases~~. ~~Having practice-oriented hydrology research in mind, w~~It~~e~~ includes~~d~~ elements underrepresented in literature,~~,~~ but crucial to the earth systems sciences: uncertainty and integrating inter-disciplinary ~~interdisciplinary ~~knowledge. We evaluated MCDA as a transdisciplinary process along the three framework steps ~~(sect. 4.3)~~. MCDA ~~can ~~only partially contributes to co-design (step 1~~; Table 4; sect. 4.3.1~~)~~, and is not appropriate for building the research team or alliances with regional partners~~. However, if ~~MCDA is ~~understood as a process ~~that ~~including ~~es ~~problem structuring, ~~it ~~MCDA ~~can ~~supports joint problem framing.~~.~~ Stakeholder analysis ~~is one method to~~helps identifying those ~~that should be involved~~to involve~~; we recommend using a relatively simple questionnaire~~. ~~MCDA p~~Problem structuring includes creativity techniques ~~also helps~~for defining shared objectives and~~objectives shared by researchers and practitioners, and we recommend creativity techniques for engaging participants in~~ designing ~~developing ~~options (i.e., ~~system configurations~~FEWS configurations). ~~Joint MCDA problem structuring allowed bringing a relatively large number of stakeholders on board. ~~The main ~~strength ~~benefit of ~~a ~~participatory MCDA ~~process ~~lies in ~~its contribution to knowledge ~~co-production (step 2~~; sect. 4.3.2~~). Interdisciplinary knowledge integration and uncertainty were rarely emphasized in literature and could be research contributions to the earth systems sciences. Both are strengths of MCDA~~ provides an integrative, consistent framework to produce transferable knowledge across different disciplines, using marginal value functions (sect. 2.2.8). Interdisciplinary integration was rarely emphasized in the reviewed literature, and could be a contribution of future research in the earth systems sciences. The same goes for explicit consideration of various types of uncertainty, which is a strength of MCDA (sect. 4.1.2)~~. MCDA ~~can ~~also provides clear results and ~~identify ~~consensus ~~system configurations~~FEWS by integrating conflicting stakeholder interests into the model~~ (sect. 3.5)~~. MCDA does not well achieve many aspects of co-dissemination (step 3). ~~An MCDA process can enable social learning of participants, but does not foster training and capacity building. Finally, ~~MCDA ~~includes discussing ~~results are discussed with stakeholders~~ and constructing improved consensus configurations~~, but this focus is narrow. ~~ (step 3, sect. 4.3.3). ~~MCDA does not achieve ~~-~~important elements such as analyzing governance mechanisms, and ~~developing or~~ implementing actions and products. In FANFAR, we thus carried out complementary activities~~ to be more comprehensive, beyond the limits of MCDA~~. ~~We received feedback that stakeholders are willing to use the FANFAR system (Figure 7), and indications that it is already effective.~~

As many others, we believe that transdisciplinary research contributes to solving ~~the ~~our global~~complex~~ problems~~ our world is facing~~, and can~~to~~ advance~~ing~~ scientific progress. We hope that this paper documents and helps to better understand ~~such ~~a transdisciplinary process in a complex setting: producing a good ~~flood forecast and alert system~~FEWS for West Africa, together with many stakeholders. We contribute to literature by analyzing the strengths ~~but also~~and limits of a comprehensive, participatory ~~-~~MCDA process for such endeavors. We encourage our colleagues from the earth system sciences to engage in transdisciplinary research with stakeholders and society.

## ~~5~~7 Data availability

The data will be available on the Eawag Research Data Institutional Collection (ERIC: https://opendata.eawag.ch), DOI: ~~<~~ xxx~~DOI Link to be provided once paper is accepted >~~

## ~~6~~8 Supplement link ~~(< will be included by Copernicus, if applicable >)~~

In the Supplementary Information (SI), we provide ample material to guide readers unfamiliar with Multi-Criteria Decision Analysis (MCDA) through all steps. This includes a Methods section (generating ~~system~~ FEWS configurations, eliciting weights, MCDA model, sensitivity analyses, stakeholder feedback), and a Results section (stakeholder analysis, objectives and attributes, ~~system configurations~~FEWS, predictions, ~~marginal~~ value functions, weights, MCDA results, and stakeholder feedback).

## ~~7~~9 Author contribution

Judit Lienert: Conceptualization, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, writing – original draft preparation, rewriting, revisions. Jafet Andersson: Funding acquisition, investigation, project administration, writing – review & editing. Daniel Hofmann: Data curation, formal analysis, visualization, validation, writing – original draft preparation. Francisco Silva Pinto: Formal analysis, investigation, methodology, project administration, writing – review & editing. Martijn Kuller: Investigation, project administration, supervision, writing – review & editing.

## ~~8~~10 Competing interests

The authors declare that they have no conflict of interest.

## ~~9~~11 Special issue statement

This paper was prepared for the special issue: “Contributions of transdisciplinary approaches to hydrology and water resources management”.~~; < The statement on a corresponding special issue will be included by Copernicus, if applicable >~~

## ~~10~~12 Acknowledgements

~~This project~~FANFAR received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 780118: FANFAR (Reinforced cooperation to provide operational flood forecasting and alerts in West Africa (FANFAR, 2021). We thank Eawag, the Swiss Federal Institute of Aquatic Science and Technology, for ~~providing additional~~ funding ~~to support the project during~~ the COVID-19 prolongation period. We sincerely thank Abdou Ali, Addi

Shuaib Olorunoje, Aishat Ibrahim, Alice Aubert, Aytor Naranjo, Berit Arheimer, Bernard Minoungou, Bode Gbobaniyi, Cletus Musa, David Gustafsson, Emilie Breviere, Emmanuel Mathot, Fabrizio Pacini, Fridolin Haag, Kevin Schönholzer, Léonard Santos, Lorna Little, Melissande Machefer, Mohamed Hamatan, Philipp Beutler, Sara Schmid, Tharcisse Ndayizigiye, and Umar Magashi for ~~their~~ invaluable support ~~when~~ of ~~preparing the~~ workshops, data collection~~ng data~~, and ~~data~~ analysis. We are truly grateful for the inspiring collaboration with ~~the~~ workshop participants, and ~~we wish to~~ sincerely thank them for their kindness and enthusiasm ~~to work with us~~. We ~~also~~ thank ~~three~~ five anonymous reviewers and the editors of this Special Issue for ~~their~~ helpful feedback.

## ~~11~~13 References

Abdullah, M. F., Siraj, S., and Hodgett, R. E.: An overview of Multi-Criteria Decision Analysis (MCDA) application in managing water-related disaster events: Analyzing 20 years of literature for flood and drought events, Water, 13, 10.3390/w13101358, 2021.

Aich, V., Liersch, S., Vetter, T., Fournet, S., Andersson, J. C. M., Calmanti, S., van Weert, F. H. A., Hattermann, F. F., and Paton, E. N.: Flood projections within the Niger River Basin under future land use and climate change, Science of the Total Environment, 562, 666-677, 10.1016/j.scitotenv.2016.04.021, 2016.

Andersson, J., Ali, A., Arheimer, B., Crochemore, L., Gbobaniyi, B., Gustafsson, D., Hamatan, M., Kuller, M., Lienert, J., Machefer, M., Magashi, U., Mathot, E., Minoungou, B., Naranjo, A., Ndayizigiye, T., Pacini, F., Silva Pinto, F., Santos, L., and Shuaib, A.: Flood forecasting and alerts in West Africa − experiences from co-developing a pre-operational system at regional scale, EGU General Assembly 2020, 4–8 May 2020, Online, https://doi.org/10.5194/egusphere-egu2020-7660, 2020a.

Andersson, J., Santos, L., Isberg, K., Gustafsson, D., Musuuza, J., Minoungou, B., and Crochemore, L.: Deliverable: D3.2. Report documenting and explaining the hydrological models, European Union, Horizon 2020, Innovation Action ICT programme, Project: 780118 FANFAR; DOI: 10.13140/RG.2.2.17369.85601 pp. 41, 2020b.

Andersson, J. C. M., Arheimer, B., Traoré, F., Gustafsson, D., and Ali, A.: Process refinements improve a hydrological model concept applied to the Niger River basin, Hydrological Processes, 31, 4540-4554, https://doi.org/10.1002/hyp.11376, 2017.

Arheimer, B., Pimentel, R., Isberg, K., Crochemore, L., Andersson, J. C. M., Hasan, A., and Pineda, L.: Global catchment modelling using World-Wide HYPE (WWH), open data, and stepwise parameter estimation, Hydrol. Earth Syst. Sci., 24, 535-559, 10.5194/hess-24-535-2020, 2020.

Arvai, J. L., Gregory, R., and McDaniels, T. L.: Testing a structured decision approach: Value-focused thinking for deliberative risk communication, Risk Analysis, 21, 1065-1076, 2001.

Basher, R.: Global early warning systems for natural hazards: systematic and people-centred, Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences, 364, 2167-2180, 10.1098/rsta.2006.1819, 2006.

Belton, V., and Stewart, T. S.: Multiple Criteria Decision Analysis: An Integrated Approach, Springer Science+Business Media Springer, Dordrecht, The Netherlands, 2002.

Berg, P., Almén, F., and Bozhinova, D.: HydroGFD3.0: a 25 km global near real-time updated precipitation and temperature data set, Earth Syst. Sci. Data Discuss., 2020, 1-18, 10.5194/essd-2020-236, 2020.

Bierens, S., Boersma, K., and van den Homberg, M. J. C.: The legitimacy, accountability, and ownership of an impact-based forecasting model in disaster governance, Politics Gov., 8, 445-455, 10.17645/pag.v8i4.3161, 2020.

Bond, S. D., Carlson, K. A., and Keeney, R. L.: Generating objectives: Can decision makers articulate what they want?, Manage. Sci., 54, 56-70, https://doi.org/10.1287/mnsc.1070.0754, 2008.

Caniglia, G., Luederitz, C., von Wirth, T., Fazey, I., Martin-Lopez, B., Hondrila, K., Konig, A., von Wehrden, H., Schapke, N. A., Laubichler, M. D., and Lang, D. J.: A pluralistic and integrated approach to action-oriented knowledge for sustainability, Nat. Sustain., 4, 93-100, 10.1038/s41893-020-00616-z, 2021.

Carr, G., Gomes, S. L., Hermans, L., Krueger, T., Lindquist, E., Thaler, T., and Slater, L.: Special issue: Contributions of transdisciplinary approaches to hydrology and water resources management Hydrol. Earth Syst. Sci., https://hess.copernicus.org/articles/special_issue1092.html, 2021.

Shiny: Web Application Framework for R R package version 1.5: https://cran.r-project.org/web/packages/shiny/index.html, 2020.

Convertino, M., Annis, A., and Nardi, F.: Information-theoretic portfolio decision model for optimal flood management, Environmental Modelling & Software, 119, 258-274, 10.1016/j.envsoft.2019.06.013, 2019.

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

de Brito, M. M., Evers, M., and Almoradie, A. D. S.: Participatory flood vulnerability assessment: a multi-criteria approach, Hydrol. Earth Syst. Sci., 22, 373-390, 10.5194/hess-22-373-2018, 2018.

Eisenführ, F., Weber, M., and Langer, T.: Rational Decision Making, Springer, Berlin, Heidelberg, New York, 2010.

Evers, M., Almoradie, A., and de Brito, M. M.: Enhancing Flood Resilience Through Collaborative Modelling and Multi-criteria Decision Analysis (MCDA), in: Urban Disaster Resilience and Security: Addressing Risks in Societies, edited by: Fekete, A., and Fiedrich, F., Urban Book Series, 221-236, 2018.

FANFAR: Operational flood forecasting and alerts in West Africa. A transdisciplinary international project funded by the European Union.: https://fanfar.eu and https://cordis.europa.eu/project/id/780118/results, access: 30.09.2021, 2021.

Figueira, J., and Roy, B.: Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure, European Journal of Operational Research, 139, 317-326, https://doi.org/10.1016/S0377-2217(01)00370-8, 2002.

Gregory, R., Failing, L., Harstone, M., Long, G., McDaniels, T., and Ohlson, D.: Structured decision making: A practical guide to environmental management choices, Wiley-Blackwell Publishing, 2012a.

Gregory, R., Long, G., Colligan, M., Geiger, J. G., and Laser, M.: When experts disagree (and better science won't help much): Using structured deliberations to support endangered species recovery planning, Journal of Environmental Management, 105, 30-43, 10.1016/j.jenvman.2012.03.001, 2012b.

Grimble, R., and Wellard, K.: Stakeholder methodologies in natural resource management: A review of principles, contexts, experiences and opportunities, Agricultural Systems, 55, 173-193, 10.1016/S0308-521X(97)00006-1, 1997.

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

Haag, F., Reichert, P., Maurer, M., and Lienert, J.: Integrating uncertainty of preferences and predictions in decision models: An application to regional wastewater planning, Journal of environmental management, 252, 109652-109652, 10.1016/j.jenvman.2019.109652, 2019b.

Haag, F., Zuercher, S., and Lienert, J.: Enhancing the elicitation of diverse decision objectives for public planning, European Journal of Operational Research, 279, 912-928, 10.1016/j.ejor.2019.06.002, 2019c.

Haag, F., Aubert, A. H., and Lienert, J.: Supporting environmental decisions with multiple objectives, multiple stakeholders, and uncertainty with the ValueDecisions web app, Under review, subm.

Howard, R. A.: Decision-analysis - Practice and promise, Manage. Sci., 34, 679-695, 10.1287/mnsc.34.6.679, 1988.

Jahn, T., Bergmann, M., and Keil, F.: Transdisciplinarity: Between mainstreaming and marginalization, Ecological Economics, 79, 1-10, 10.1016/j.ecolecon.2012.04.017, 2012.

Janis, I. L.: Victims of Groupthink, Houghton Mifflin, New York, 1972.

Keeney, R. L.: Decision-Analysis - an Overview, Oper. Res., 30, 803-838, https://doi.org/10.1287/opre.30.5.803, 1982.

Keeney, R. L.: Value-focused thinking: Identifying decision opportunities and creating alternatives, European Journal of Operational Research, 92, 537-549, https://doi.org/10.1016/0377-2217(96)00004-5, 1996.

Kendall, M. G.: A new measure of rank correlation, Biometrika, 30, 81-93, 10.2307/2332226, 1938.

Kuller, M., Silva Pinto, F., Schönholzer, K., and Lienert, J.: Effective risk communication for early flood warning in West-Africa, EGU General Assembly 2020, Due to coronavirus pandemic, EGU2020 has gone virtual with Sharing Geoscience Online; https://doi.org/10.5194/egusphere-egu2020-9425, 4.-8.05.2020, 2020.

Kuller, M., Schönholzer, K., and Lienert, J.: Creating effective flood warnings: a framework from a critical review, J. Hydrol., 602, 126708, https://doi.org/10.1016/j.jhydrol.2021.126708, 2021.

Kuller, M., Beutler, P., and Lienert, J.: Preference change in stakeholder group-decision processes in the public sector: extent, causes and implications, in prep.

Lang, D. J., Wiek, A., Bergmann, M., Stauffacher, M., Martens, P., Moll, P., Swilling, M., and Thomas, C., J.: Transdisciplinary research in sustainability science: practice, principles, and challenges, Sustainability Science, 7 25-43, https://link.springer.com/article/10.1007/s11625-011-0149-x, 2012.

Lemos, M. C., and Morehouse, B. J.: The co-production of science and policy in integrated climate assessments, Glob. Environ. Change-Human Policy Dimens., 15, 57-68, 10.1016/j.gloenvcha.2004.09.004, 2005.

Lienert, J., Schnetzer, F., and Ingold, K.: Stakeholder analysis combined with social network analysis provides fine-grained insights into water infrastructure planning processes, Journal of Environmental Management, 125, 134-148, 10.1016/j.jenvman.2013.03.052, 2013.

Lienert, J., Scholten, L., Egger, C., and Maurer, M.: Structured decision-making for sustainable water infrastructure planning and four future scenarios, EURO Journal on Decision Processes, 3, 107-140, http://dx.doi.org/10.1007/s40070-014-0030-0, 2015.

Lienert, J., Andersson, J., Hofmann, D., Silva Pinto, F., and Kuller, M.: Deliverable D2.2. Report on the co-design workshops in FANFAR to create a flood forecast and alert system for West Africa European Union, Horizon 2020, Innovation Action ICT programme, Project: 780118 FANFAR, Persistent URL: https://www.dora.lib4ri.ch/eawag/islandora/object/eawag:21545, Deliverable D2.2, 29, 2020.

Liu, S., Maclean, K., and Robinson, C.: A cost-effective framework to prioritise stakeholder participation options, Euro Journal on Decision Processes, 7, 221-241, 10.1007/s40070-019-00103-7, 2019.

Lumbroso, D., Brown, E., and Ranger, N.: Stakeholders' perceptions of the overall effectiveness of early warning systems and risk assessments for weather-related hazards in Africa, the Caribbean and South Asia, Nat. Hazards, 84, 2121-2144, 10.1007/s11069-016-2537-0, 2016.

Marttunen, M., and Hamalainen, R. P.: The decision analysis interview approach in the collaborative management of a large regulated water course, Environmental Management, 42, 1026-1042, 10.1007/s00267-008-9200-9, 2008.

Marttunen, M., Lienert, J., and Belton, V.: Structuring problems for Multi-Criteria Decision Analysis in practice: A literature review of method combinations, European Journal of Operational Research, 263, 1-17, 10.1016/j.ejor.2017.04.041, 2017.

Marttunen, M., Belton, V., and Lienert, J.: Are objectives hierarchy related biases observed in practice? A meta-analysis of environmental and energy applications of Multi-Criteria Decision Analysis, European Journal of Operational Research, 265, 178-194, 10.1016/j.ejor.2017.02.038, 2018.

Marttunen, M., Haag, F., Belton, V., Mustajoki, J., and Lienert, J.: Methods to inform the development of concise objectives hierarchies in multi-criteria decision analysis, European Journal of Operational Research, 277, 604-620, 10.1016/j.ejor.2019.02.039, 2019.

Massazza, G., Tarchiani, V., Andersson, J. C. M., Ali, A., Ibrahim, M. H., Pezzoli, A., De Filippis, T., Rocchi, L., Minoungou, B., Gustafsson, D., and Rosso, M.: Downscaling Regional Hydrological Forecast for Operational Use in Local Early Warning: HYPE Models in the Sirba River, Water, 12, 3504, https://doi.org/10.3390/w12123504 2020.

Mauser, W., Klepper, G., Rice, M., Schmalzbauer, B. S., Hackmann, H., Leemans, R., and Moore, H.: Transdisciplinary global change research: the co-creation of knowledge for sustainability, Current Opinion in Environmental Sustainability, 5, 420-431, 10.1016/j.cosust.2013.07.001, 2013.

Montibeller, G., and von Winterfeldt, D.: Cognitive and motivational biases in decision and risk analysis, Risk Analysis, 35, 1230-1251, 10.1111/risa.12360, 2015.

Nachappa, T. G., Piralilou, S. T., Gholamnia, K., Ghorbanzadeh, O., Rahmati, O., and Blaschke, T.: Flood susceptibility mapping with machine learning, multi-criteria decision analysis and ensemble using Dempster Shafer Theory, J. Hydrol., 590, 10.1016/j.jhydrol.2020.125275, 2020.

NBA (Niger Basin Authority): SATH-NBA (Satellite based water monitoring and flow forecasting for the Niger River Basin: http://www.sath.abn.ne/, access: 09.11.2020, 2020.

Nka, B. N., Oudin, L., Karambiri, H., Paturel, J. E., and Ribstein, P.: Trends in floods in West Africa: analysis based on 11 catchments in the region, Hydrol. Earth Syst. Sci., 19, 4707-4719, 10.5194/hess-19-4707-2015, 2015.

Norstrom, A. V., Cvitanovic, C., Lof, M. F., West, S., Wyborn, C., Balvanera, P., Bednarek, A. T., Bennett, E. M., Biggs, R., de Bremond, A., Campbell, B. M., Canadell, J. G., Carpenter, S. R., Folke, C., Fulton, E. A., Gaffney, O., Gelcich, S., Jouffray, J. B., Leach, M., Le Tissier, M., Martin-Lopez, B., Louder, E., Loutre, M. F., Meadow, A. M., Nagendra, H., Payne, D., Peterson, G. D., Reyers, B., Scholes, R., Speranza, C. I., Spierenburg, M., Stafford-Smith, M., Tengo, M., van der Hel, S., van

Putten, I., and Osterblom, H.: Principles for knowledge co-production in sustainability research, Nat. Sustain., 3, 182-190, 10.1038/s41893-019-0448-2, 2020.

O'Hagan, A.: Expert knowledge elicitation: subjective but scientific, American Statistician, 73, 69-81, 10.1080/00031305.2018.1518265, 2019.

OCHA: West and Central Africa: flooding situation, UN Office for the Coordination of Humanitarian Affairs (OCHA); https://www.unocha.org/, 2, 2020.

Pang, B., Yue, J. J., Huang, Z. Q., and Zhang, R.: Parameter uncertainty assessment of a flood forecasting model using multiple objectives, Journal of Flood Risk Management, 12, 10.1111/jfr3.12493, 2019.

Passerotti, G., Massazza, G., Pezzoli, A., Bigi, V., Zsótér, E., and Rosso, M.: Hydrological Model Application in the Sirba River: Early Warning System and GloFAS Improvements, Water, 12, 620, 10.3390/w12030620, 2020.

Paulus, P. B., and Yang, H. C.: Idea generation in groups: A basis for creativity in organizations, Organizational Behavior and Human Decision Processes, 82, 76-87, 10.1006/obhd.2000.2888, 2000.

Perera, D., Seidou, O., Agnihotri, J., Rasmy, M., Smakhtin, V., Coulibaly, P., and Mehmood, H.: Flood early warning systems: a review of benefits, challenges and prospects, United Nations University Institute for Water, Environment and Health (UNU-INWEH); https://inweh.unu.edu/, Hamilton, Canada, 26, 2019.

Pham, B. T., Luu, C., Phong, T. V., Nguyen, H. D., Le, H. V., Tran, T. Q., Ta, H. T., and Prakash, I.: Flood risk assessment using hybrid artificial intelligence models integrated with multi-criteria decision analysis in Quang Nam Province, Vietnam, J. Hydrol., 592, 10.1016/j.jhydrol.2020.125815, 2021.

Pictet, J., and Bollinger, D.: Extended use of the cards procedure as a simple elicitation technique for MAVT. Application to public procurement in Switzerland, European Journal of Operational Research, 185, 1300-1307, 10.1016/j.ejor.2006.05.051, 2008.

R Core Team: R: A language and environment for statistical computing, Version 3.3.1, R Foundation for Statistical Computing, http://www.R-project.org/, Vienna, Austria, 2018.

Reed, M. S., Graves, A., Dandy, N., Posthumus, H., Hubacek, K., Morris, J., Prell, C., Quinn, C. H., and Stringer, L. C.: Who's in and why? A typology of stakeholder analysis methods for natural resource management, Journal of Environmental Management, 90, 1933-1949, 10.1016/j.jenvman.2009.01.001, 2009.

Reichert, P., Schuwirth, N., and Langhans, S. D.: Constructing, evaluating and visualizing value and utility functions for decision support, Environmental Modelling & Software, 46, 283-291, https://doi.org/10.1016/j.envsoft.2013.01.017, 2013.

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

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

Ronco, P., Bullo, M., Torresan, S., Critto, A., Olschewski, R., Zappa, M., and Marcomini, A.: KULTURisk regional risk assessment methodology for water-related natural hazards - Part 2: Application to the Zurich case study, Hydrol. Earth Syst. Sci., 19, 1561-1576, 10.5194/hess-19-1561-2015, 2015.

Rosenhead, J., and Mingers, J.: Rational Analysis for a Problematic World Revisited: Problem Structuring Methods for Complexity, Uncertainty and Conflict, John Wiley & Sons, Chichester, 2001.

Roudier, P., Ducharne, A., and Feyen, L.: Climate change impacts on runoff in West Africa: a review, Hydrol. Earth Syst. Sci., 18, 2789-2801, 10.5194/hess-18-2789-2014, 2014.

Samanta, S., Koloa, C., Pal, D. K., and Palsamanta, B.: Flood Risk Analysis in Lower Part of Markham River Based on Multi-Criteria Decision Approach (MCDA), Hydrology, 3, 10.3390/hydrology3030029, 2016.

Schneider, F., Giger, M., Harari, N., Moser, S., Oberlack, C., Providoli, I., Schmid, L., Tribaldos, T., and Zimmermann, A.: Transdisciplinary co-production of knowledge and sustainability transformations: Three generic mechanisms of impact generation, Environmental Science & Policy, 102, 26-35, 10.1016/j.envsci.2019.08.017, 2019.

Silva Pinto, F., and Lienert, J.: Report activities to establish co-design committee, stakeholder analysis, European Union, Horizon 2020, Innovation Action ICT programme; Project: 780118 FANFAR, Deliverable D2.1, 48, 2018.

Sultan, B., Lejeune, Q., Menke, I., Maskell, G., Lee, K., Noblet, M., Sy, I., and Roudier, P.: Current needs for climate services in West Africa: Results from two stakeholder surveys, Clim. Serv., 18, 10.1016/j.cliser.2020.100166, 2020.

Tang, Z. Q., Zhang, H., Yi, S. Z., and Xiao, Y. F.: Assessment of flood susceptible areas using spatially explicit, probabilistic multi-criteria decision analysis, J. Hydrol., 558, 144-158, 10.1016/j.jhydrol.2018.01.033, 2018.

UNISDR: Early warning practices can save lives: good practices and lessons learned, United Nations Secretariat of the International Strategy for Disaster Reduction (UNISDR); https://www.undrr.org/, Bonn, Germany, 67, 2010.

WMO: 2020 State of climate services: risk information and early warning systems, WMO (World Meteorological Organization), 2020.

WMO (World Meteorological Organization): Regional Climate Outlook Forums: https://public.wmo.int/en/our-mandate/climate/regional-climate-outlook-products, access: 25.02.2021, 2021.

Wuelser, G., Adler, C., Breu, T., Hadorn, G. H., Wiesmann, U., and Pohl, C.: On which common ground to build? Transferable knowledge across cases in transdisciplinary sustainability research, Sustainability Science, 16, 1891–1905, 10.1007/s11625-021-01010-0, 2021.

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