Weighting factors for LCA-a new set from a global survey

Subasi, A.B.; Askham, C.; Sandorf, E.D.; Dias, L.C.; Campbell, D.; Tas, E.F.; ... ; Cinelli, M.

The International Journal of Life Cycle Assessment
https://doi.org/10.1007/s11367-024-02330-w

LCI METHODOLOGY AND DATABASES

Weighting factors for LCA—a new set from a global survey

Ayşe Bayazıt Subaşı[1] · Cecilia Askham[2] · Erlend Dancke Sandorf[3] · Luis Cândido Dias[4] · Danny Campbell[5] · Elçin Filiz Taş[1] · Norihiro Itsubo[6] · Christine Betty Nagawa[7] · Christine Mugumya Kyarimpa[8] · Mamadou Djerma[9] · Bazoin Sylvain Raoul Bazie[9,10] · Marco Cinelli[11,12]

Received: 1 September 2023 / Accepted: 21 May 2024

Abstract

Purpose This paper provides global weights (weighting factors) for the three endpoint impact categories (areas of protection (AoPs)) of the United Nations Environment Programme (UNEP) Life Cycle Initiative's "Global Guidance for Life Cycle Impact Assessment Indicators and Methods" (GLAM) project, namely human health, ecosystem quality, and natural resources and ecosystem services.

Methods A discrete choice experiment (DCE) was conducted to elicit the preferences of respondents on the GLAM AoPs, and they were then used to calculate the respective weights. Responses were obtained from a subset of countries pertaining to each income level defined by the World Bank (i.e. low, lower-middle, upper-middle, and high). The adimensional (between 0 and 1) weights were derived using two different approaches: econometric and multiple criteria decision analysis (MCDA). The econometric approach obtained weights by transforming the estimated preference parameters from a multinomial logit model. The MCDA approach obtained weights representing the vectors that best reconstitute the choices of each individual, using linear programming to fit an additive value function.

Results When considering responses from all income groups, the weights from the econometric approach are 0.42, 0.31, and 0.26 for human health, ecosystem quality, and natural resources and ecosystem services, respectively. Following the same order for the AoPs, the weights from the MCDA approach are 0.41, 0.32, and 0.27. For high-income countries, ecosystem quality has the highest weight; for upper-middle-income countries, ecosystem quality and human health have the same weights using the econometric approach, while in the MCDA approach, human health is weighted higher than ecosystem quality. For the two lower income country groups, the priority is given to human health with both approaches. Recommendations for the use of these weights are also provided, as well as a comparison with other existing weights.

Conclusion The two methods obtained similar weights overall, although with some differences when disaggregated by income groups. The weights proposed in this paper are suitable for decision-makers or users who want to use survey-derived weights for endpoint-based LCA when using the AoPs of GLAM. These weights can be used in projects where the decision-makers do not want to or have no resources to identify a set of weights themselves, or when decision-makers are not involved.

Keywords Life cycle assessment · Life cycle impact assessment · Endpoint · Weighting · Area of protection · UNEP GLAM

Abbreviations

AoP	Area of protection
CF	Characterization factor
DALY	Disability-adjusted life year
DCE	Discrete choice experiment
EI99	Eco-indicator 99
EINES	Expected increase in number of extinct species
ESM	Electronic supplementary material
EQ	Ecosystem quality
Eq.	Equation
GDP	Gross domestic product
GLAM	Global Guidance for Life Cycle Impact Assessment Indicators and Methods
HH	Human health
IN	Impact after normalization
IO	Impact on the original endpoint scale
Inv	Inventory
ISO	International Organization for Standardization

Communicated by Matthias Finkbeiner.

Extended author information available on the last page of the article

Published online: 19 August 2024

IUCN	International Union for Conservation of Nature
LCA	Life cycle assessment
LCIA	Life cycle impact assessment
LCIn	Life Cycle Initiative
LIME	Life-cycle impact assessment method based on endpoint modelling
MCDA	Multiple criteria decision analysis
MNL	Multinomial logit model
NRandES	Natural resources and ecosystem services
NUTS	Nomenclature of Territorial Units for Statistics
NV	Normalization value
UNEP	United Nations Environment Programme
USD	United States dollar
WEMSS	Weighting Methods Selection Software
WF	Weighting factor

1 Introduction

Life cycle assessment (LCA) uses multiple categories to express the level of impact that the systems under study cause (Dias et al. 2019). These impacts can be of two different types in LCA, namely midpoints (e.g. climate change, ozone depletion) and endpoints (e.g. human health, ecosystem quality). When these impacts are calculated in the impact assessment phase, weighting can be part of the interpretation phase of an LCA. According to ISO14044, weighting is an optional element in LCA, and it is described as "converting and possibly aggregating indicator results across impact categories using numerical factors based on value-choices" (ISO 2006). Even though it is optional, weighting can be helpful for decision-makers during the interpretation of the LCIA results. Weighting in LCA can be performed with two main objectives (Itsubo 2015): (i) to identify the most important impact categories and thereafter the life cycle stages that contribute to these impacts and (ii) to understand which system performs comprehensively better than the others, usually via a single score. Sala and Cerutti (2018) classified weighting methods for LCA into five main groups: (1) single item (based on, e.g. physical properties), (2) distance-to-target (based on, e.g. policy targets or planetary boundaries), (3) panel-based (e.g. based on surveys of experts, citizens, or government panels), (4) monetary valuation (based on monetary estimation from, e.g. observed preferences, revealed preferences, stated preferences, budget constraints, abatement cost, damage cost), and (5) meta-models (based on multiple weighting factors from the combination of other weighting sets). Lippiat (2007) and Huppes et al. (2006) used panel methods at the midpoint level, while LIME3 (Itsubo et al. 2018), Ecoindicator99 (EI99) (Goedkoop and Spriensma 2001), and ReCiPe (Huijbregts et al. 2017) used damage cost as weighting at the endpoint level. The distance-to-target weighting was used by Castellani et al. (2016) for Europe, and recently (e.g. Bjorn et al. (2020) and Sala et al. (2020)) related to the planetary boundaries identified by Rockström et al. (2009) and Steffen et al. (2015).

To achieve the above-mentioned objectives of weighting in LCA, some form of value judgement is needed to make the impact categories comparable, incorporate the preferences of the affected stakeholders, and possibly aggregate the impact categories in a single or subset of single scores. Some of the most prominent approaches for eliciting stakeholders' preferences are described below. One approach to elicit preferences involves having respondents make choices between competing alternatives, e.g. holiday packages, cars, and planned motorways, where each alternative is described by its attributes. Researchers can use this choice data to estimate people's preferences for these alternatives and how much weight they put on each of the attributes, e.g. how much weight they put on cost, environmental impact, or social benefit. Methods to elicit preferences in this way are broadly classified as stated preference methods because respondents state their preferences through the choices they make in a contingent or hypothetical market. This paper uses a specific type of stated preference method called a discrete choice experiment (DCE).[1] Stated preference methods have been used since the mid-1970s to understand people's preferences for, among other things, ecosystem management policies, environmental protection, environmentally certified products, energy demand, recreation, industrial projects, and policy development (Alriksson and Öberg 2008; Rakotonarivo et al. 2016; Hoyos 2010), as well as farmers' attitude toward agricultural policies (Mamine et al. 2020), consumer's attitude toward food labelling (Lombardi et al. 2017), and triggering factors for eco-innovations (Horbach et al. 2012). Both DCEs and other types of conjoint analysis have been used to derive weights for impact categories in LCA. For example, Bai et al. (2018) use a conjoint analysis approach where they have a sample of stakeholders rank alternatives from most to least preferred and infer preferences based on these rankings to assign importance to four midpoint indicators. The LIME methodology (Itsubo et al. 2018), on the other hand, uses a DCE to elicit preferences, which are then used to derive weights. The latest iteration of the LIME methodology, LIME3 (Murakami et al. 2018), develops weighting factors for four LCA endpoints (human health, social assets, biodiversity, and primary production),

[1] In some parts of the literature, this method, among many others, is classified as conjoint analysis. This paper uses the term DCE to be precise with respect to the method used and the fact that this is firmly grounded in random utility theory (Louivere et al. 2010).

for the G20, G8, and a subset of other specific countries (see Itsubo et al. 2018 for details).

Another approach for eliciting (and also structuring) preferences is multiple criteria decision analysis (MCDA). MCDA is a methodology that has been developed to support complex decision-making. It provides a structured process to formulate a decision-making problem by identifying the alternatives to be evaluated (if not available yet) and the criteria to assess them. In addition, it provides tools to shape an evaluation model that can be used to aggregate the performance of each alternative on each criterion and the preferences of the stakeholders to provide a final decision recommendation. The latter can be one of these four types (Cinelli et al. 2020):

1. Rank alternatives from the best to the worst
2. Classify alternatives in preference-ordered classes
3. Choose the most preferred alternative(s) according to the predefined constraints
4. Cluster alternatives according to some similarity features

Preferences in MCDA are elicited during the problem structuring phase, as well as during the development of the evaluation model. These can include, among others, the importance of the criteria, the interactions between the criteria, different forms of thresholds, and the level of compensation among the criteria (Dias et al. 2018). In MCDA, direct and indirect approaches can be used to elicit the preferences (Hüllermeier and Slowinski, 2024). In the case of direct elicitation, the parameters of the model are defined directly by the decision-maker. It has been demonstrated that this approach can be challenging for the respondent, as they struggle to understand the information they are asked to provide (e.g. trade-offs between the criteria, level of interaction between the criteria), or they might not have the time to dedicate to lengthy exercises that are associated with them. In order to overcome these issues, indirect preference elicitation approaches have received a lot of attention in MCDA (Doumpos and Zopounidis 2011). They require the respondent to provide (relatively) easy local or holistic judgements (e.g. sort, choice, ranking) on a set of (well-known) reference alternatives (Cinelli et al. 2022a, b) (i.e. decision examples). These judgements are then used to infer the preference model of the respondent. MCDA methods that use indirect preferences are also called disaggregation methods, and they have experienced a notable increase in popularity in many application areas, due to the easiness of provision of these local or holistic judgements (Hüllermeier and Slowinski 2024). This is also the case for the environmental studies. Some examples include the synthesis of nanomaterials (Cinelli et al. 2019; Kadzinski et al. 2020), energy accidents (Cinelli et al. 2019), land use suitability (Androulaki and Psarras 2016), climate change communication (Zerva et al. 2021), transportation (Dias et al. 2021), and resource management (Tervonen et al. 2015; Zheng and Lienert 2018). As far as the weighting of the impact categories in LCA is concerned, to the best of the authors' knowledge, there are no previous applications of MCDA methods based on indirect preferences.

1.1 Main contributions of the paper

The "Global Guidance for Life Cycle Impact Assessment Indicators and Methods" (GLAM) project is supported by the Life Cycle Initiative (LCIn), hosted by the United Nations Environment Programme (UNEP) (LCIn 2023). It aims at establishing consensus on life cycle impact assessment (LCIA) indicators and methods. One of the objectives of this project is to provide at least one set of weighting factors for the three AoPs that the GLAM methodology is based on. However, previous literature does not provide weighting sets that fit the AoPs of the GLAM methodology. To fill this gap, the weighting subtask of the GLAM project developed these weights. DCE and an MCDA method based on indirect preferences were used for this purpose as they fit with the requirements of the GLAM project, as well as the available resources in the research team. Using DCE and an MCDA disaggregation method, this paper presents the results from a large-scale study eliciting preferences from over 3000 citizens from different countries around the world, and it answers these three main research questions:

- Which weights for GLAM AoPs could future users apply, representing the preferences of a large sample of the world population?
- How do these weights vary across country income level groups?
- How can the weights be used in the GLAM LCIA methodology?

This research also presents three contributions that are applicable to the LCA practice more broadly than to the sole GLAM project. Firstly, it provides a conceptualization of LCA, environmental science, and economics in relation to weighting and aggregation. Secondly, it provides an implementation strategy for the application of DCE and MCDA disaggregation in future LCA studies, including respondents from low-income countries too (which is the first demonstration of its type in LCA). Lastly, this work addresses an objective of a more methodological nature, which is to assess to which extent using two rather different approaches on the same set of preferences leads to similar results.

In this paper, Section 2 presents the methodology developed to select and apply the suitable methods within the GLAM project. Section 3 presents the results and discusses the weights obtained from both methods, including a comparison of the weights presented in this paper and existing weights from other approaches. It also covers when the weights derived from this global survey can be used and how they fit within the LCA domain and beyond. Section 4 discusses the results, and Section 5 concludes the paper.

2 Methodology

The methodology employed by the weighting subtask started with a selection of a set of suitable weighting methods for the GLAM project (Section 2.1). After this selection, an implementation strategy for the chosen methods was devised (Section 2.2).

2.1 Selection of suitable methods for the GLAM project

The LCIn, hosted by UNEP, is a public–private, multi-stakeholder partnership enabling the global use of life cycle knowledge by public and private sector stakeholders, providing a global forum for science-based, consensus-building processes to support sustainability (UNEP 2023). LCIn initiated the GLAM project to establish consensus on life cycle impact assessment (LCIA) indicators and methods. In phase 1 and 2 of GLAM, guidance on indicators such as greenhouse gas emissions (climate change), particulate matter impact on human health, water use related impacts, land use impacts on biodiversity, acidification, eutrophication, human toxicity, natural resources, ecotoxicity, land use impacts on soil quality, and cross-cutting issues were provided. Phase 3 started in 2020 and aims to establish a comprehensive, consistent, and global LCIA method, including normalization and weighting (UNEP-GLAM 2021).

Phase 3 of the GLAM project has four main task forces: (1) human health, (2) ecosystem quality, (3) natural resources and ecosystem services, and (4) normalization, weighting, and cross-cutting issues. The weighting subtask consisted of a group of LCA and decision-making experts (for the list of members of this subtask, see Annex A in Electronic Supplementary Material-ESM who operated under task force 4. Their aim was to develop guidelines to select the most relevant methodologies to elicit weights and to build a consistent set of weights covering all AoPs of the GLAM LCIA methodology (UNEP LCI 2021).

Through the history of LCIA, different weighting approaches have been developed and applied to ease communication of the results to decision-makers. The choice of the weighting method influences the communication of the results, and LCA/decision analysts can find it difficult to choose which is the most relevant. In order to facilitate this choice, the weighting subtask developed the WEighting Methods Selection Software (WEMSS) (Cinelli et al. 2023) which is a freely available software (access link: https://mcda.cs.put.poznan.pl/wemss/index.php) for decision-makers to select a weighting method according to their needs. The simple interface leads the user through a question-and-answer process which reduces the number of suitable methods according to the choices made for each question. In the WEMSS, 35 weighting methods are assessed according to 50 key decision-making features. These features were selected after a consensual process within the weighting subtask. They were then used to assess each weighting method by two or three members of the weighting subtask. The resulting assessments were then discussed within the subtask and then finalized by a consensus procedure (Cinelli et al. 2022a, b).

The process to identify the suitable weighting methods for the GLAM project involved multiple rounds of interaction among the members of the weighting subtask, as well as some members from the GLAM project management. These iterative sessions led to the agreement of the answers to the WEMSS as reported in Fig. 1.

Some remarks related to the agreed answers:

1. Independence from the set of systems being evaluated: weights should be independent from the set of systems being evaluated. *Rationale*: the weights should be applicable to any system that the analyst would be interested to work with.
2. Scientific validity: published in a peer-reviewed journal or book. *Rationale*: the chosen method should be recognized by the scientific community.
3. Method transparency: method algorithms and value choices can be partly explained. *Rationale*: the need for transparency is important for understanding the calculation strategies of the weights. The complexity of the methods should be acknowledged, though the maximum effort should be employed to explain how the chosen algorithms work and the influence of the modelling choices on the results.
4. Uncertainty characterization: the uncertainties should be characterized, but stochastic characterization is not required. *Rationale*: from the viewpoint of users, it was acknowledged that the information about the range (e.g. max and min) and the average/median would be sufficient.
5. Area of protection metrics: weights should be directly related or relatable to the GLAM AoP metrics. *Rationale*: the foreseen application of the weights is in an additive aggregation model, which requires the weights to be linked to the GLAM AoP metrics to define the conversion coefficients (trade-offs).
6. Geographical resolution: no geographical differentiation. *Rationale*: the GLAM methodology aims to be applicable on the global scale. One could argue that the weights (at least those derived from surveys) should be linked

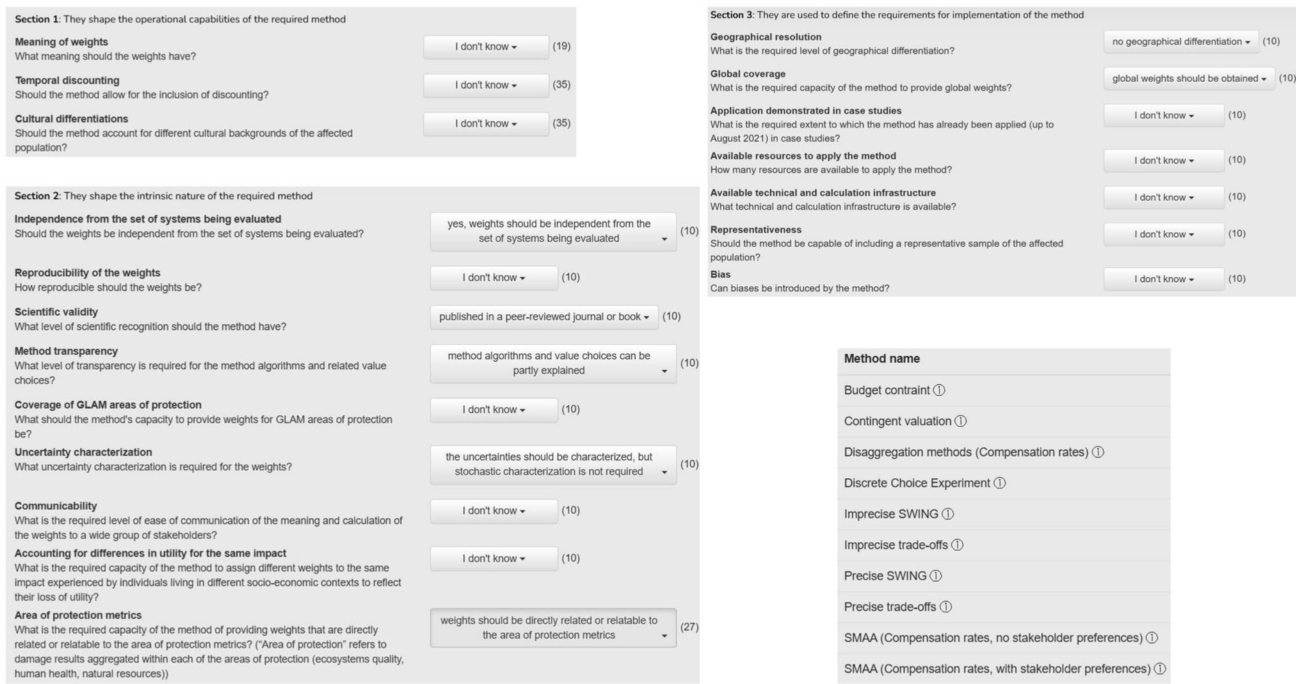

Fig. 1 Answers agreed for the selection of the weighting methods for the GLAM project and list of suitable methods recommended by the WEMSS (bottom right)

to the impact locations (e.g. people suffering from the impacts) and not impact sources, which may be in different regions. This is often not possible in LCA (geographical resolution of inventory and impacts are often lost); thus, a region-specific weighting is not needed for the GLAM project.

7. Global coverage: global weights should be obtained. *Rationale*: the weights provided should be covering the impacts on a global scale.
8. Only the features for which priority was agreed by the subtask received attention and were entered as requirements. The remaining ones were left blank (i.e. "I do not know" answer, meaning no option is excluded).

The agreed answers to the WEMSS led to a shortlist of 10 suitable weighting methods for the GLAM project, shown in the bottom right of Fig. 1. To finally choose one or a subset of methods to be implemented for the development of the weights, the subtask decided to invite all the subtask members to express their availability for applying one or more of the shortlisted methods within the remaining timeframe of the project. A DCE was the method that had the most members available to apply it. It is also important to recall that DCE has been successfully applied for endpoint weighting in LCA previously in LIME (Itsubo et al. 2004, 2012; Murakami et al. 2018). This also increased the trust in the method for this application area. In addition, the weighting subtask organized a dedicated workshop with experts (Danny Campbell and Erlend Dancke Sandorf) in DCE on 14 January 2022 to further discuss the suitability of DCE to the GLAM project objectives. After this session, it was clear that the application of DCE would be demanding but manageable within the timeframe of the subtask, and so DCE was selected to calculate the weights within the GLAM project.

Another weighting method (i.e. a disaggregation method providing weights for a multi-attribute additive value model) out of the suitable set of 10 candidates for the GLAM project was also chosen to calculate the weights. This disaggregation method belongs to the MCDA domain. This choice was driven by the fact that the survey used to apply the DCE provided input data that could also be used by the MCDA disaggregation approach, and this type of expertise was available in the subtask team. The weighting subtask agreed that using two methods to derive these weights provided a valuable opportunity for robustness analysis.

Overall, the weighting subtask group was involved in the DCE work through giving input to the survey development and providing their suggestions in meetings where the work was presented and discussed. Thus, the weighting subtask was involved in a collaborative process for many of the steps along the way to developing the weighting factors presented in this paper. However, it should be noted that they were not asked to vote about the set of binding requirements for the methodology or whether they agree that weights calculated from this work are the final set of weights that should be used in the GLAM methodology. There has thus not been a

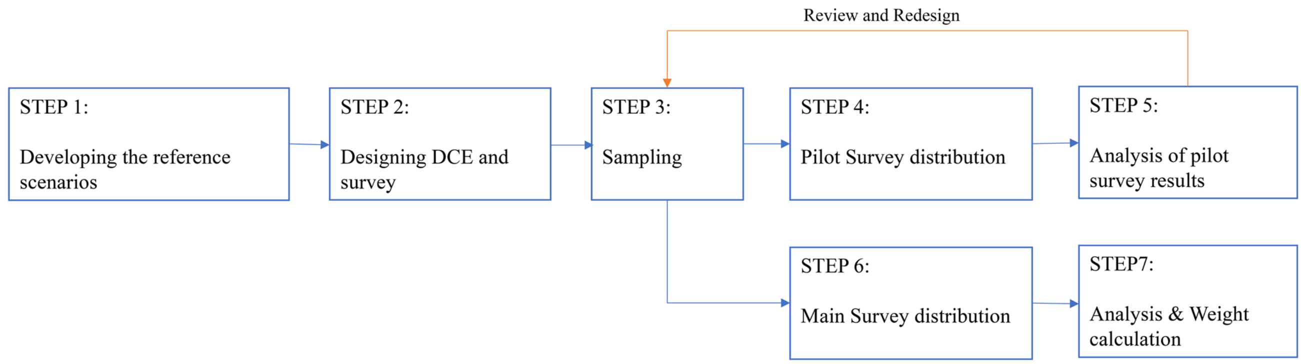

Fig. 2 Procedure for the development of the survey and the calculation of the weights

minority statement process. Instead, the members that had reservations on the chosen methods were encouraged to provide recommendations to tackle them (see Section 4.5).

2.2 Application of the weighting methods for the GLAM project

A multi-step process was developed by the weighting subtask to calculate the weights as part of the GLAM project. This included the development of a survey to elicit the preferences of a target population regarding the three AoPs defined in the GLAM LCIA framework: human health (HH), ecosystem quality (EQ), and natural resources and ecosystem services (NRandES). Endpoint, rather than midpoint modelling, has been an active choice in the GLAM methodology. It has been previously argued that uncertainty increases when moving from midpoint to endpoint and that in some cases, midpoint indicators are more desirable (Bare et al. 2000). However, Verones et al. (2017) argue that "modelling beyond the midpoint introduces relevant additional information and hence that the midpoint result is less environmentally relevant than the damage result". Since midpoint results stop within the cause-effect chain, the environmental consequence on human health or ecosystem quality is often unclear. In addition, many models, especially for ecosystem quality, tend to model directly to the endpoint (e.g. Scherer et al. (2023) and Pierrat et al. (2023)) and do not have a natural midpoint indicator. Thus, it can be argued that using endpoint modelling is now at least as certain as midpoint modelling. This is particularly relevant where there is no midpoint indicator. An example of a situation where midpoint is deemed more certain than endpoint is climate change, where the damage can be more uncertain (due to factors like socio-economic responses) than the radiative forcing potential. The inclusion of more mechanisms of damage, as well as their consequences for humans and the environment, can be described as increased comprehensiveness or relevance, rather than uncertainty. Inclusion of more effects means that it is possible to achieve a more comprehensive assessment of damage, which is not necessarily less uncertain.

Figure 2 summarizes the methodology developed for the preparation of the survey, its design, dissemination, and weight calculation (full versions of the pilot and main survey are provided in Annex B and C in the ESM, respectively).

The starting point (step 1) consisted of defining the "reference scenarios", which represent the current level of damage to the different AoPs. These are directly related to the normalization values (NVs) in LCIA. In the current study, final NVs for the GLAM-LCIA methodology could not be used, as they were not yet available (calculating these was a parallel task within the GLAM project). Thus, estimates for damages to the three AoPs were derived based on previous normalization work, including those in previous GLAM phases (UNEP LCI 2021). In step 2, the number of attributes (i.e. AoPs) and performance levels, choice cards, graphics, and other parts of the survey were designed. In step 3, the type of respondents, sample size, and the survey administration method were defined. Qualtrics software (Qualtrics XM 2023) was used to implement the survey. In step 4, the pilot survey was made available in English and Turkish and was distributed between November and December 2022. The dissemination took place via the subtask and their networks (snowballing), as well as face-to-face surveys carried out in Uganda. The results from the pilot survey were analysed in step 5 and led to the revision of the survey. The analysis of the feedback and suggestions from the pilot survey are given in detail in Annex D of ESM. This input was incorporated in the main survey, which was distributed between February and

April 2023 (step 6 in Fig. 2). In step 7, the data was analysed, and weights were calculated as described in Section 3.1 below. In addition, the results were interpreted, reported, and disseminated.

2.3 Discrete choice experiment design

2.3.1 Development of the reference scenarios

The weighting subtask interfaced with each of the AoP task forces of the GLAM project, which provided the NVs used to shape the reference scenarios. For HH AoP, two NVs were used to explore the sensitivity to the absolute values stated in the reference scenarios, corresponding to 19 days[2] and 55 days[3] of healthy life lost per person per year in the world due to disability or early death.

For EQ AoP, the NV was expressed as 12% of terrestrial species at risk of extinction. Only land occupation impacts on biodiversity were considered environmental stressors for this estimate due to the lack of the other methods (characterization factors) available in GLAM at the time of this study. Terrestrial species included in the calculations were mammals, amphibians, birds, reptiles, and vascular plants. The risk of extinction was defined by the International Union for Conservation of Nature Red List of Threatened Species (IUCN 2023).

For NRandES, the values for natural resources and ecosystem services were calculated separately and then summed. For natural resources (NR), AoP task force updated the characterization factors of natural resource use from the future welfare loss model by Huppertz et al. (2019). Together with these CFs for fossil fuels, minerals, and energy carriers, they used the production amount of these natural resources for the year 2015, which is retrieved from USGS mineral commodity summaries (U.S. Geological Survey 2022) and BP Statistical Review of World Energy (BP 2022). The valuations for ecosystem services (ES) have been developed by the task force and applied using four ecosystem service pathways (erosion resistance, filtration, groundwater generation, and biotic production) described by De Laurentiis et al. (2019) and Bos et al. (2016). These values are not yet published and under preparation for publication. Based on this work for the NRandES AoP, the annual losses amounted to 6,480 billion USD (reference year 2018) globally, which is equivalent to 7.5% of global gross domestic product (GDP) based on data retrieved from the World Data Bank (World Bank 2023a). In order to make these numbers relatable to the general population in a given country, this was adjusted to per capita GDP by calculating 7.5% of the average per capita GDP for the different income groups for 2018 (World Bank 2023b) as shown in Table 1.

Table 1 Normalization references of NR and ES AoP for different income levels

Income level (defined by World Bank)	Loss in USD
High	3398
Upper-middle	711
Lower-middle	170
Low	51

2.3.2 Attributes, levels, and choice task design

In the DCE, respondents were asked to choose between three scenarios described by the three AoPs. Each AoP could take one of the nine different levels shown in Table 2 as positive and negative deviations from the reference value. The levels used spanned the range from a 100% reduction to a doubling of the impact of the AoP. The first alternative was always the reference alternative and was characterized by no deviation from the reference values (see Section 2.3.1). To generate the hypothetical scenarios 1 and 2, an efficient experimental design was used (Scarpa and Rose 2008). Each set of three scenarios forms a choice task. Respondents are asked to choose their preferred alternative among the three presented in each choice task. Through their choices, respondents reveal their preferences. An example of a choice task is shown in Fig. 3.

In our context, a reduction in the impact of an attribute (AoP) is preferable to an increase in the impact. Therefore, to generate the initial design, it is assumed that, ceteris paribus, a reduction in impact is preferred to an increase in impact, and a small negative prior (−0.0001) was specified for all attributes. The prior is simply the assumed impact of the value of each AoP to the overall value of the alternative when generating the design and reflects the researcher's best knowledge about the direction and magnitude of this impact. The practical implication of a small negative prior is that people prefer reductions in impact over increases, but the magnitude of this effect is not clear, but the design now considers this when combining attribute levels into choice tasks.

The initial (and updated) design was a D_p-efficient design optimized for the multinomial logit model (Scarpa and Rose 2008) using Ngene (Rose et al. 2018). The initial design for the pilot survey comprising nine choice tasks is shown in Table 3. For instance, in choice task number 1, HH is set to 100% more than the reference

[2] HH damage derived from the global burden of disease (reference year 2019) linked to both environmental and occupational health was 18.7 days/person per year (https://www.healthdata.org/research-analysis/gbd).

[3] Fifty-five days of healthy life (1.05E+09 DALY/year) was used as one of the reference values, obtained from using LC-Impact method to calculate the human health impacts for all effects associated with global emissions data.

Table 2 Attributes and levels used in the choice card design in the survey

Attributes (AoPs)	Levels
HH	– 100%, – 75%, – 50%, – 25%, 0% (reference), + 25%, + 50%, + 75%, + 100%
EQ	– 100%, – 75%, – 50%, – 25%, 0% (reference), + 25%, + 50%, + 75%, + 100%
NRandES	– 100%, – 75%, – 50%, – 25%, 0% (reference), + 25%, + 50%, + 75%, + 100%

scenario value in scenario 1, and it is set to 100% less than the reference scenario (i.e. zero) in scenario 2. EQ in choice card number 1 is set to 75% more than the reference scenario value in scenario 1 and to 100% less than the reference scenario in scenario 2. NRandES in choice card 1 is set to 25% less than the reference scenario value in scenario 1 and to 25% more than the reference scenario value for scenario 2.

The results of the pilot study were used to update the design of the main survey. There was no evidence to suggest updating the priors nor using different priors for the different country groups. The main change included the need to test the two different levels for the HH AoP reference values. This meant that the updated design was larger, allowing for more parameters, and second-order interaction terms between the attributes. Zero priors are assumed for the interaction terms when generating the design. The updated design comprised 27 choice tasks of three alternatives, each described by the same three attributes (AoPs). The 27 choice tasks were allocated into three blocks of nine choice tasks each and respondents were randomly allocated to one of the blocks when entering the survey. The design for the main survey is shown in Table 4 (note “ct” is a choice task).

2.4 Sampling strategy

The target respondents for the survey were ordinary citizens, since the GLAM project is focused on the preferences of the general population, rather than on experts.

Fig. 3 An example of a choice card shown to respondents

Table 3 Level configuration of nine choice cards in pilot survey

Choice card no.	Reference scenario			Scenario 1			Scenario 2		
	HH	EQ	NRandES	HH	EQ	NRandES	HH	EQ	NRandES
1	0	0	0	1	0.75	−0.25	−1	−1	0.25
2	0	0	0	0.75	−1	−1	−0.75	1	1
3	0	0	0	−0.75	1	−1	1	−0.75	1
4	0	0	0	−1	−1	0.25	1	1	−0.75
5	0	0	0	−1	1	0.75	0.75	−1	−1
6	0	0	0	1	−1	1	−1	0.5	−0.75
7	0	0	0	0.75	0.5	−1	−1	−1	0.75
8	0	0	0	0.5	−0.75	−1	−0.75	1	1
9	0	0	0	0.75	−0.75	0.75	−1	0.75	−1

Given the global focus of the GLAM project, as well as the financial and time constraints, the income levels defined by the World Bank (World Bank 2023c) (i.e. low, lower-middle, upper-middle, and high) were chosen as a framework to target a subset of countries from each income group. In terms of sample size, based on previous experience in the application of a similar survey with the LIME3 methodology (Itsubo et al. 2018), and accounting for practical constraints (budget, time, and human resources), the subtask aimed for 500 responses per income level, with a total of at least 2000 respondents.

The sampling strategy involved selecting a few countries within each income group and sampling from those countries. In the study, multiple sampling strategies were applied.

In Uganda, where the survey was carried out face-to-face in the field, a combination of purposive and random sampling was implemented. More specifically, five districts (Bulisa, Rakai, Masaka, Jinja, and Busia) were

Table 4 Efficient design used in the main survey

		Reference scenario			Scenario 1			Scenario 2		
		HH	EQ	NRandES	HH	EQ	NRandES	HH	EQ	NRandES
Block 1	ct_2	0	0	0	1	−1	−0.75	−0.75	1	−0.75
	ct_10	0	0	0	−1	0.25	−0.75	0.75	−1	1
	ct_11	0	0	0	0.25	1	−0.25	−1	−1	1
	ct_12	0	0	0	0.5	0.75	−0.5	1	−1	−1
	ct_16	0	0	0	1	−0.75	−0.75	−1	1	−0.75
	ct_18	0	0	0	−0.25	−1	1	−1	1	1
	ct_20	0	0	0	1	−0.25	0.75	1	1	−1
	ct_26	0	0	0	0.75	0.75	−1	−1	−1	0.75
	ct_27	0	0	0	−1	1	0.75	1	0.75	−0.25
Block 2	ct_4	0	0	0	−0.75	−1	0.75	−1	1	1
	ct_5	0	0	0	−1	1	−1	1	−0.25	−1
	ct_8	0	0	0	−1	0.25	1	1	−1	0.75
	ct_9	0	0	0	0.75	1	−1	−1	1	1
	ct_14	0	0	0	1	0.75	−1	1	−0.5	1
	ct_15	0	0	0	0.25	−1	−1	−0.25	1	1
	ct_17	0	0	0	0.5	1	−0.75	−1	−1	1
	ct_19	0	0	0	−1	1	−0.75	1	0.75	−0.75
	ct_21	0	0	0	−1	−0.25	1	1	−1	−0.5
Block 3	ct_1	0	0	0	1	1	−1	1	−1	1
	ct_3	0	0	0	0.25	−1	−1	−0.75	−1	1
	ct_6	0	0	0	−1	1	1	0.25	−1	−1
	ct_7	0	0	0	1	−0.25	1	−0.75	1	−0.75
	ct_13	0	0	0	−0.75	1	−1	−0.25	0.75	1
	ct_22	0	0	0	−1	0.75	0.75	−1	1	−1
	ct_23	0	0	0	0.75	1	−1	1	−1	1
	ct_24	0	0	0	1	−1	−1	0.25	0.75	−1
	ct_25	0	0	0	0.75	−1	1	−0.75	1	−1

targeted based on purposive sampling. Five villages were included per district, and five households were selected at random in each chosen village. In Burkina Faso, where the survey was also administered face-to-face in the field, sampling took place in eight of the 13 regions of the country: the Centre region, the Central Plateau region, the Cascades region, the Hauts-Bassins region, the Sud Ouest region, the Centre-Est region, the Centre-Sud region, and the Centre-Ouest region. One household was selected at random in the urban centre of each region, and then, every 10th household was sampled from that starting point.

For Japan (Tokyo), China (Shanghai), and India (Mumbai), a survey company was used to recruit participants. The survey company used quota sampling, based on age and gender and the need to provide a sample size of 500 respondents for each country/city. The targeted respondents were people who lived in the chosen cities and belonged to an ad hoc mailing list gathered by the survey company. Respondents were randomly asked to participate in the survey. The quota sampling approach for this survey company excluded respondents over the age of 60, under 18, with no education, and only primary education.

Another survey company was also involved in carrying out the survey face-to-face in the field with 250 respondents in Türkiye (Turkey) (more specifically Adana, Ankara, Bursa, Erzurum, Gaziantep, İzmir, Kayseri, Malatya, Samsun, Balıkesir, Trabzon), according to statistical regions of Türkiye (Eurostat 2023). The rest of the respondents from Türkiye, mainly from Istanbul, were obtained via the snowballing strategy.

A snowball sampling approach was used for all other respondents.

The survey was administered in two ways, (i) autonomously filled in via web-based access to the survey and (ii) mediated by enumerators via face-to-face interviews. The first option was adopted for lower-middle, upper-middle, and high-income countries, where access to the internet is high or at least substantial. According to the World Bank, the percentage of the population using the internet is 89% in high-income countries, 76% in upper-middle-income countries, and 45% in lower-middle-income countries (World Bank 2023d). This contrasts with low-income countries, where only 19% of the population uses the internet. Consequently, the second type of administration (face-to-face interviews in the field) was adopted for low-income countries. It was also used partially for Türkiye (upper-middle-income). The enumerators used tablets or phones to access the Qualtrics survey and directly filled in the responses during the interview.

2.5 Survey design

Qualtrics software was used to create the survey, which consisted of five sections: (i) consent, (ii) introduction, (iii) selection of preferred scenarios, (iv) demographic data, and (v) feedback.

In the first section, the consent page allowed the participants to understand what the survey was about and decide whether they were willing to participate (ethics committee approvals can be found in Annex E in the ESM). If the respondents chose not to give consent, they were led directly to the end of the survey.

In the second section of the survey, the aim of the survey was briefly explained, and background information on the three AoPs, including visual aids to enhance understanding as shown for the EQ AoP in Fig. 4. The complete pilot and main surveys are available in the ESM, Annexes B and C, respectively.

After this, the survey instructions were provided to the respondents explaining what was expected from them, including how to navigate through the survey and make their choices.

In the third section of the survey, preferences were elicited from respondents by asking them to choose their preferred scenario in each choice task. Nine choice tasks were shown to each respondent both in the pilot and main surveys.

In the fourth section of the survey, demographic data was collected, including age distribution in 5-year increments (above the age of 18), gender, type of area of residence (either urban or not), and level of education. The demographic distribution of respondents of the main survey is given in Annex F in the ESM.

The last section included feedback on the survey, and it differed between the pilot and the main survey. In the pilot survey, suggestions for improvement and feedback were collected and documented (see Annex D in the ESM for the pilot survey and Annex G for the main survey). Since the feedback from respondents was essential for the improvement of the survey design for the main survey, the pilot survey had more open-ended questions. These questions aimed at eliciting the respondents' opinions about the comprehensibility of the task, the visual aids, text explanations (introduction section), choice cards, and any further suggestions. In the main survey, there was only one question asking if they had any suggestions to improve the survey, allowing them to give feedback. Moreover, there was an optional e-mail field in the survey, so that respondents that wished to be informed about the results of the study could be contacted.

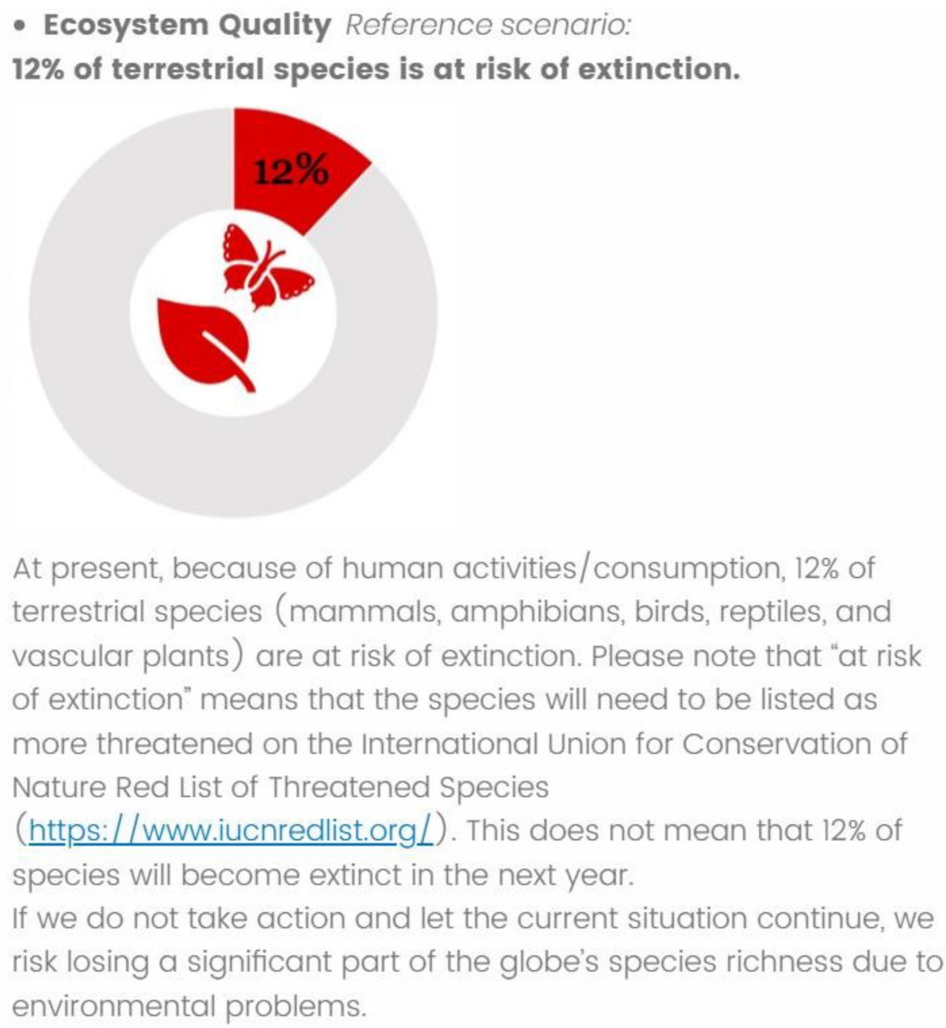

Fig. 4 The visual aid developed for EQ AoP

2.6 Survey distribution and redesign

The pilot survey was distributed between November and December 2022. Dissemination was via the subtask and their networks (snowballing), as well as face-to-face surveys carried out in Uganda. As described above, the results from the pilot survey were analysed and led to the revision of the survey (see Annex D in the ESM). The following changes were made for the main survey:

- The aim of the study was emphasized in the introduction section.
- The text description of the attributes (AoPs) was revised in the choice tasks to include specification of the units of impacts; the survey was made available in more languages. The number of languages increased from only in Turkish and English for the pilot survey, to 11 languages in total for the main survey (Chinese, English, French, German, Italian, Japanese, Norwegian, Portuguese, Spanish Turkish, and Urdu).
- The normalization subtask provided a second NV for the HH, which led to a refinement of the survey design.

The survey company completed their main survey distribution for Japan, China, and India between the 10th and 22nd of February 2023. The snowballing distribution work involved the members of the subtask distributing information about the survey to their networks and contacts. Those distributing the information about the survey also requested that the people receiving the information would also recruit friends and family with diverse interests and demographics. This snowballing action took place from the 10th of March until the 15th of April 2023. The face-to-face surveys in Uganda, Burkina Faso (low-income), and Türkiye were conducted in the same time period.

Table 5 Example of alternatives shown to one respondent and corresponding preferred alternative (i.e. a_x in choice column)

Choice task	Reference scenario (a_1)			Alternativea_2			Alternativea_3			Choice
	HH	EQ	NRandES	HH	EQ	NRandES	HH	EQ	NRandES	
1	0	0	0	0.75	1	−1	−1	1	1	a_1
2	0	0	0	0.25	−1	−1	−0.25	1	1	a_2
3	0	0	0	0.5	1	−0.75	−1	−1	1	a_3
…	…	…	…	…	…	…	…	…	…	…
9	0	0	0	−1	0.25	1	1	−1	0.75	a_1

2.7 Weights calculation approaches

Weights obtained for the GLAM LCIA methodology fit with the weighting endpoint type 1 as defined by Itsubo (2015) (more details on the use of the new set of weights are provided in Section 4.3). They were calculated using two different approaches. The first one is based on random utility theory, and it is named "econometric" approach. The second one uses deterministic value theory, and it is named "MCDA disaggregation" approach.

As inputs, the choices of the surveyed individuals are used (see an example in Table 5).

2.7.1 Weights calculation with the econometric approach

The econometric approach for the weights is grounded in random utility theory (Manski 1977; McFadden 1974). The weights are obtained by transforming the parameters of the estimated utility function. It is assumed that respondents, when asked to choose between alternatives described by the AoPs, choose the alternative that gives them the highest utility, i.e. they maximize utility. Let the utility of individual n choosing alternative i in choice task t be described by Equation (Eq.) 1 and the linear-in-the-parameters random utility function (Manski 1977):

$$U_{nit} = s\sum_{k=1}^{K}\beta_k X_{knit} + \varepsilon_{nit} \tag{1}$$

where β_k is a parameter to be estimated for the kth AoP, X_{knit} is the AoP level, s is a scale parameter that is inversely related to the variance of error term, and ε_{nit} a type I extreme value distributed error term capturing everything not explained by the AoPs and observable characteristics of the respondent included in the model. Under standard assumptions, the choice probabilities can be obtained using the multinomial logit model (MNL) (Ben-Akiva and Steven 1985; McFadden 1974) in Eq. 2:

$$Pr\left(y_n|\beta_k, X_{nit}\right) = \prod_{t=1}^{T}\frac{exp(s\sum_{k=1}^{K}\beta_k X_{knit})}{\sum_{j=1}^{J}exp\left(s\sum_{k=1}^{K}\beta_k X_{knjt}\right)} \tag{2}$$

where y_n is the sequence of choices made by individual n. Observed heterogeneity arising from differences in how HH was displayed was accounted for (either as 19 or 55 days in the reference scenario) using interaction terms ($D^{55} = 1$ if in the 55 days' reference treatment and 0 otherwise). The following linear (random) utility function is estimated, where HH, EQ, and NRandES replace X, and the subscript k is suppressed to clearly show the AoPs in Eq. 3:

$$U_{nit} = \beta_{i0} + \beta_1 HH_{nit} + \beta_2 EQ_{nit} + \beta_3 NR\&ES_{nit} + \beta_4 HH_{nit}D^{55} + \varepsilon_{nit} \tag{3}$$

where β_{i0} is an alternative specific constant for alternative i ($J-1$ constants are estimated, where J is the total number of alternatives), β_1, β_2, and β_3 are the estimated preference intensities for each of the three AoPs, and β_4 is the estimated marginal effect of being in the HH display group "55 days". When the model is estimated, no constraints are placed on the values that β_0, β_1, β_2, β_3, and β_4 can take, and a larger absolute value of β_{1-4} means that if an alternative has a larger reduction in the AoP impact relative to the other alternatives, it is more likely to be chosen, all else equal, and if it has a larger increase in the impact of the AoP, it is less likely to be chosen, all else equal (conditional on the parameters being estimated with a negative sign).

In models run on the entire dataset, or pooled datasets for country groups, it is necessary to control for relative unobserved differences (Swait and Louviere 1993). For example, there are likely unobserved differences between countries, which must be considered. To account for them, the following relative scale specification is used in Eq. 2:

$$s = \frac{\sum_{c=1}^{C}s_c I_c}{s_1 I_1} \tag{4}$$

where s_c is the relative scale parameter for country c and I_c is an indicator equal to 1 if the respondent is in country c and 0 otherwise. To normalize the relative scales, the scale of country 1 is set to unity.

The weights obtained from this random utility model are a rescaling of the estimated preference intensities (β_1, β_2, and β_3) of the model such that they add up to one.

For AoP k, the weight can be calculated as in Eq. 5:

Table 6 Examples of constraints to be respected in the MCDA disaggregation approach based on the choices made by the respondent

Question	Chosen alternative	Implied comparisons	Implied constraints
1	a_1	$I(a_1) < I(a_2)$ $I(a_1) < I(a_3)$	$0 \leq 0.75w_1 + 1w_2 - 1w_3$ $0 \leq -1w_1 + 1w_2 + 1w_3$
2	a_2	$I(a_2) < I(a_1)$ $I(a_2) < I(a_3)$	$0.25w_1 - 1w_2 - 1w_3 \leq 0$ $0.25w_1 - 1w_2 - 1w_3 \leq -0.25w_1 + 1w_2 + 1w_3$
3	a_3	$I(a_3) < I(a_1)$ $I(a_3) < I(a_2)$	$-1w_1 - 1w_2 + 1w_3 \leq 0$ $-1w_1 - 1w_2 + 1w_3 \leq 0.5w_1 + 1w_2 - 0.75w_3$
…	…	…	…
9	a_1	$I(a_1) < I(a_2)$ $I(a_1) < I(a_3)$	$0 \leq -1w_1 + 0.25w_2 + 1w_3$ $0 \leq 1w_1 - 1w_2 + 0.75w_3$

$$w_k = \frac{exp\left(-\frac{\partial U}{\partial X_k}\right)}{\sum_{k=1}^{K} exp\left(-\frac{\partial U}{\partial X_k}\right)} \tag{5}$$

Standard errors for the weights are obtained using the delta method, which is then used to obtain the 95% confidence intervals.

2.7.2 Weights calculation with the MCDA disaggregation approach

The MCDA disaggregation approach for the weights is grounded in deterministic value theory, developed since the 1970s (Dyer and Sarin 1979; Keeney 1992; Keeney and Raiffa 1979; Krantz et al. 1971). The weights obtained with this approach represent the vectors that best reconstitute the nine choices of each individual, assuming that the alternatives are chosen solely on the basis of the performance of the three AoPs.

This approach assumes an additive linear model for each respondent (Eq. 6), which implies that the overall impact (the less, the better) of an alternative a_j is a weighted sum of the impacts on the three endpoints:

$$I(a_j) = \sum_{k=1}^{K} w_k I_k(a_j) \tag{6}$$

where $K = 3$ and indices 1, 2, and 3 represent the three AoPs HH, EQ, and NRandES, respectively.

The answer to each question would satisfy two constraints if the respondent fully agreed with a given model, stating the chosen alternative has less overall impact than the other two. An example is presented in Table 6.

Each respondent originates a set of 18 pairs (a_x, a_y) such that $I(a_x) < I(a_y)$ according to the respondent's nine answers. Since the corresponding system of inequalities might be infeasible, an error term e_c, to be minimized, can be added to each constraint, replacing the constraint $I(a_x) < (a_y)$ by $I(a_x) < I(a_y) + e_c$ ($c = 1,\ldots,18$).

Using a linear programming formulation (Dias et al. 2021) derived from the well-known MCDA disaggregation approach (Jacquet-Lagrèze and Siskos 2001; Matsatsinis et al. 2018; Siskos et al. 2005), one can find the weights vector (w_1, w_2, w_3) that minimizes the maximum error term or, in an alternative formulation, the weights vector that minimizes the sum of the error terms (the latter approach was adopted here). Such formulation finds a vector of weights that respects all the constraints or, if unable to do that, provides the weights that minimize the penalty function (in this case, the sum of the errors).

When it is possible to find a vector of weights that respects all the constraints, this might be just one such vector among many other vectors that would also respect all the constraints. Therefore, a post-optimization stage (Dias et al. 2021) can optimize a secondary benefit function among the set of weights that respect the 18 constraints of the respondent. In this work, the secondary objective is to maximize the minimum weight, which avoids null or near null weights for the endpoints, if possible. In particular, if the vector of equal weights respects all the constraints, then this would be the chosen vector. The post-optimization strategy is also followed in the cases where the set of constraints is

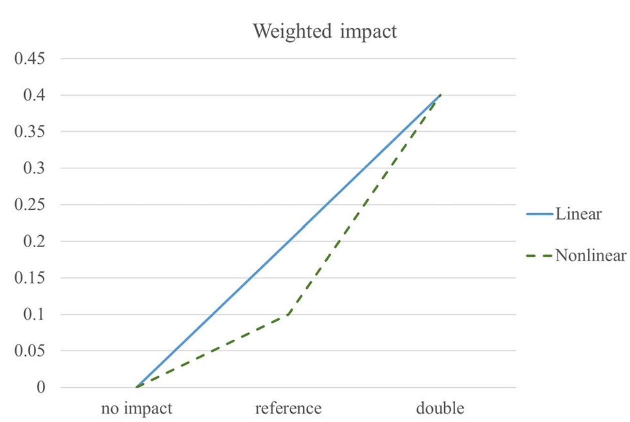

Fig. 5 Example of piece-wise linear value function in relation to a weighted impact

infeasible, by optimizing the secondary objective under the condition that the original optimal penalty value found is not worsened.

The MCDA disaggregation approach has been developed to account for two types of value functions for the respondents. The first type is the linear value function (see straight line in Fig. 5 as an example), which assumes that a change in impact is valued the same, independently of the starting point and direction (i.e. either above or below the reference). The second type is the nonlinear value function, more precisely the piece-wise linear, where the same change in impact can be valued differently depending on the starting value. The present work admits that the impact value differences that lie below the reference are not valued the same as the impact value differences better than the reference, e.g. losses loom larger than gains (see dashed line in Fig. 5 as an example). In this case, if $s_k^{<ref}$ and $s_k^{>ref}$ denote the weighted function slope for impacts below and above the reference $\left(a_1\right)$, respectively, for AoP k, then $w_k I_k\left(a_j\right)$ in Eq. 6 can be rewritten as $min\left\{I_k\left(a_j\right), I_k\left(a_1\right)\right\} \times s_k^{<ref} + max\left\{I_k\left(a_j\right) - I_k\left(a_1\right), 0\right\} \times s_k^{>ref}$. Also, the total weight for AoP k will be $s_k^{<ref} + s_k^{>ref}$.

The logic used to calculate the weights with piece-wise linear functions is then the same as the one used when assuming linear functions. The added flexibility of dropping the assumption of a proportional (linear) impact increases the number of cases where the model respects all the constraints, as usually occurs when more freedom is allowed, although at risk of overfitting the data (Dias et al. 2022).

Two models have thus been developed based on the MCDA disaggregation approach:

- MCDA model 1: linear value functions minimizing the error sum, followed by post-optimization
- MCDA model 2: piece-wise linear value functions minimizing the error sum, followed by post-optimization

Using one of these models yields one vector of weights (or a pair of slopes, in the piece-wise linear case) per respondent. It is then possible to make a statistical description of the responses, which in this case includes the computation of the mean weights from all of the responses, the mean weights according to each income level, as well as distinguishing the value of health in the reference scenario ("high" or "low").

3 Results

A total of 3198 respondents completed at least one choice task. The demographic information gathered from respondents to the survey is compared with that of the world in order to study how representative the respondents to the survey were for the world (Table 7). The respondents were largely between the ages of 25 and 44. The education levels were reasonably aligned with the world averages, e.g. 27% having primary and secondary school education, compared to the world average of 26%. However, the survey sample tertiary education level (bachelor + master + doctorate) was 54%, compared to the world average of 40%. More specifically, respondents with a PhD are overrepresented in the study, as they total 7.54%. The only available statistic for the population with PhD is for

Table 7 Demographics of the GLAM survey and comparison with the same statistics for the world

	This study %	World %
Age		
18–24	3.88	*7.58
25–29	16.17	7.43
30–34	18.64	7.59
35–39	16.85	7.12
40–44	17.07	6.41
45–49	9.16	5.95
50–54	7.91	5.73
55–59	6.60	5.14
60–64	2.25	4.08
65 and above	1.47	9.81
Education		
No education	6.72	10
Primary	11.32	**24
Secondary	10.04	***26
High	17.48	
Bachelor	34.77	****40
Master	12.13	
Doctorate	7.54	
Gender		
Male	49.16	50.26
Female	50.19	49.74
Non-binary/third gender	0.44	-
Prefer to self-describe	0.00	-
Prefer not to say	0.22	-
Area of living		
Urban	67.32	56.91
Suburban	18.67	-
Rural	14.01	43.09

The data for the world is gathered from World Bank (World Bank 2023e) for age and from Stata for education (Statista 2023)

*20–24 ages

**Primary

***Secondary

****Tertiary

Table 8 Weights grouped by HH reference for all income groups (econometric models)

Econometric models	Human health	Ecosystem quality	Natural resources and ecosystem services
HH reference = high	0.45 [0.43, 0.47]	0.30 [0.29, 0.31]	0.25 [0.23, 0.27]
HH reference = low	0.40 [0.39, 0.41]	0.33 [0.31, 0.34]	0.28 [0.26, 0.29]

Ninety-five percent (95%) confidence intervals within brackets

OECD countries, which shows that only about 1% of their population has a doctorate. The sample is representative of gender (49% male and 50% female, with 1% non-binary or not stated), while the comparative world statistics give a ratio of men and women of almost 50:50. The urban population is higher in the sample than for the world (67% and 57%, respectively). This is not surprising, as the survey company used for Japan, China, and India only targeted respondents from large cities (Tokyo, Shanghai, and Mumbai). The sample in this weighting survey has fewer respondents living in rural areas than the world average (14 as opposed to 43). The underrepresented groups in the survey presented in this paper are citizens above the age of 65 and those living in rural areas. More information on the distribution of respondents in different income level groups (low, lower-middle, upper-middle, and high) can be found in Annex F in the ESM.

The weights are presented below for the two approaches developed for the GLAM project (econometric and MCDA). The key findings are included in this section; the interested reader can find all models used, all the weights calculated, and the associated variability ranges in Annex H, I, and J in the ESM.

3.1 The new sets of weights

3.1.1 Weights from the econometric approach

This section reports the main findings from the application of the econometric approach. They are based on 3198 respondents (respondents that completed at least one choice task), distributed as follows:

- 795 respondents for the high-income group
- 950 respondents for the upper-middle-income group
- 513 respondents for the lower-middle-income group
- 940 respondents for the low-income group

The weights for the three AoPs for all the income groups are presented in Table 8, grouped by HH reference. In both cases, HH is the AoP with the highest weight, irrespective of the HH reference, with a weight of 0.45 and 0.40 depending on the HH reference. A consistent trend of preferences for the weights obtained with both HH references can be seen for the remaining AoPs, with EQ always in the 2nd position, and NRandES in the 3rd one.

Table 9 Weights clustered by income group and HH reference (econometric models)

Econometric models		Human health	Ecosystem quality	Natural resources and ecosystem services
All income groups	**HH reference = high + low**	**0.42 [0.41, 0.43]**	**0.31 [0.30, 0.32]**	**0.26 [0.25, 0.28]**
High-income group	**HH reference = high + low**	**0.34 [0.32, 0.36]**	**0.41 [0.40, 0.43]**	**0.25 [0.23, 0.27]**
	HH reference = high	0.38 [0.36, 0.41]	0.39 [0.37, 0.41]	0.23 [0.22, 0.24]
	HH reference = low	0.30 [0.28, 0.32]	0.44 [0.42, 0.46]	0.26 [0.25, 0.27]
Upper-middle-income group	**HH reference = high + low**	**0.36 [0.35, 0.38]**	**0.36 [0.35, 0.37]**	**0.28 [0.27, 0.29]**
	HH reference = high	0.37 [0.35, 0.39]	0.35 [0.34, 0.37]	0.28 [0.26, 0.29]
	HH reference = low	0.35[0.33, 0.37]	0.36 [0.35, 0.38]	0.28 [0.27, 0.30]
Lower-middle-income group	**HH reference = high + low**	**0.36 [0.35, 0.38]**	**0.32 [0.30, 0.33]**	**0.32 [0.31, 0.34]**
	HH reference = high	0.37 [0.35, 0.40]	0.31 [0.29, 0.33]	0.32 [0.30, 0.33]
	HH reference = low	0.35 [0.33, 0.37]	0.32 [0.30, 0.34]	0.33 [0.31, 0.34]
Low-income group	**HH reference = high + low**	**0.54 [0.51, 0.56]**	**0.24 [0.23, 0.26]**	**0.22 [0.20, 0.24]**
	HH reference = high	0.58 [0.55, 0.61]	0.22 [0.20, 0.24]	0.20 [0.18, 0.22]
	HH reference = low	0.49 [0.47, 0.52]	0.26 [0.25, 0.28]	0.24 [0.22, 0.26]

Ninety-five percent (95%) confidence intervals within brackets

Looking at the weights based on the high HH reference, there is a substantial gap of 0.15 and 0.20, respectively, between HH (0.45) and EQ (0.30) and HH and NRandES (0.25). In the case of the low HH reference, the weight gap between the first two AoP drops by 0.07, with EQ increasing its weight to 0.33. NRandES also increases its weight by 0.03, reaching 0.28.

Table 9 shows the weights grouped by income and HH reference. Weights obtained by combining the results from the scenarios with both HH references (bold font in Table 9) are shown first. The weights that consider responses for all of the income groups show a clear and statistically significant trend. HH is in the first position (0.42), EQ in the second one (0.31), and NRandES in the third position (0.26).

Moving to the high-income group and pooling both HH reference results, EQ AoP shows the highest weight of 0.41, followed by the HH AoP (0.34), and lastly by the NRandES AoP (0.25). This preference order for the AoPs is unique among the income groups: the HH AoP has the highest weight for the lower-middle (0.36) and low-income (0.54) groups, while it has the same weight (0.36) as EQ AoP for the upper-middle-income group.

The EQ AoP receives the second highest weight only for the low-income group (0.24), with NRandES following closely with a weight of 0.22. In the case of the lower-middle-income group, these (EQ and NRandES) AoPs receive the same weight (0.32).

The weights within each income group based on the different HH references provide further insights in the preferences of the respondents (italics font in Table 9). For the high-income group, irrespective of the HH reference, EQ receives the highest weight, HH the second highest, and NRandES the lowest. However, the differences between the weights are substantial when the HH reference changes from high to low. With the high HH reference, the difference between the weights for the EQ AoP and HH AoP is only 0.01; this increases to 0.14 when the HH reference is low. Looking at the other income groups, the weights assigned to HH are always higher when the high HH reference is used. The HH AoP is always the one with the highest weight, except for the upper-middle-income group with low HH reference. The low-income group stands out because it shows very high weights for both HH references, with a value of 0.58 for the high HH reference, and 0.49 for the low HH reference. Another notable finding is that, for the low-income group, the weight for this (HH) AoP with the low HH reference (0.49) is higher than the highest weight for HH in any of the other income groups, irrespective of the HH reference.

Focusing on the upper-middle and lower-middle income groups, the weights for the HH AoP are the same with high (0.37) and low (0.35) HH references. However, there are differences in the weights of the remaining AoPs. The weight of the EQ AoP is very close to the weight for the HH AoP for the upper-middle-income countries, while it is lower in the lower-middle-income countries (by 0.06 and 0.03 for high and low HH references, respectively). Lastly, in the lower-middle-income group, independent of the HH reference used, NRandES has a higher weight (0.32 and 0.33) compared to EQ (0.31 and 0.32), though the confidence intervals overlap to a large extent; hence, statistical significance is not confirmed.

3.1.2 Weights from the MCDA disaggregation approach

This section reports the main findings from the application of the MCDA disaggregation approach. They are based on 3003 replies (accounting only for the respondents that completed all the nine choice tasks), which are distributed as follows:

- 760 respondents for the high-income group
- 830 respondents for the upper-middle-income group
- 504 respondents for the lower-middle-income group
- 909 respondents for the low-income group

Based on the choices of the surveyed individuals, and depending on the model, different results are derived.

The mean weights for the three AoPs for all the income groups are presented in Table 10, grouped by HH reference. The weight of the HH AoP for all income groups combined is 0.43 for the high HH reference, while it is 0.39 in the case of low HH reference. Similarly to the econometric model, the HH AoP has always the highest weight, followed by EQ (second), and NRandES (third). In this model, this means that decreasing HH impacts by p% would be more important than decreasing the EQ impacts by p%, and reducing the NRandES impacts by p% would be the least important, for any percentage p.

For the scenarios with the high HH reference, the weight for EQ is 0.31, while the weight for NRandES is 0.26.

Table 10 Mean weights grouped by HH reference for all income groups (MCDA linear model)

MCDA linear	Human health	Ecosystem quality	Natural resources and ecosystem services
HH reference = high	0.43 [0.42, 0.44]	0.31 [0.30, 0.32]	0.26 [0.25, 0.27]
HH reference = low	0.39 [0.38, 0.40]	0.34 [0.33, 0.35]	0.27 [0.26, 0.28]

Ninety-five percent (95%) confidence intervals within brackets

Table 11 Mean weights clustered by income group and HH reference (MCDA linear model)

MCDA linear		**Human health**	**Ecosystem quality**	**Natural resources and ecosystem services**
All income groups	**HH reference = high + low**	**0.41 [0.40, 0.42]**	**0.32 [0.32, 0.33]**	**0.27 [0.26, 0.27]**
High-income group	**HH reference = high + low**	**0.36 [0.34, 0.37]**	**0.39 [0.37, 0.4]**	**0.26 [0.24, 0.27]**
	HH reference = high	0.39 [0.38, 0.41]	0.36 [0.34, 0.38]	0.25 [0.23, 0.26]
	HH reference = low	0.32 [0.30, 0.34]	0.41 [0.39, 0.43]	0.27 [0.25, 0.28]
Upper-middle income group	**HH reference = high + low**	**0.39 [0.38, 0.40]**	**0.33 [0.32, 0.34]**	**0.28 [0.26, 0.29]**
	HH reference = high	0.41 [0.39, 0.42]	0.32 [0.30, 0.34]	0.27 [0.26, 0.29]
	HH reference = low	0.38 [0.36, 0.39]	0.34 [0.33, 0.36]	0.28 [0.26, 0.30]
Lower-middle income group	**HH reference = high + low**	**0.39 [0.38, 0.40]**	**0.31 [0.29, 0.32]**	**0.31 [0.29, 0.32]**
	HH reference = high	0.39 [0.37, 0.41]	0.31 [0.29, 0.33]	0.30 [0.28, 0.32]
	HH reference = low	0.39 [0.37, 0.40]	0.30 [0.28, 0.32]	0.31 [0.29, 0.33]
Low-income group	**HH reference = high + low**	**0.48 [0.47, 0.49]**	**0.27 [0.26, 0.29]**	**0.25 [0.23, 0.26]**
	HH reference = high	0.51 [0.49, 0.53]	0.25 [0.24, 0.27]	0.24 [0.22, 0.26]
	HH reference = low	0.46 [0.43, 0.47]	0.29 [0.28, 0.31]	0.25 [0.23, 0.27]

Ninety-five percent (95%) confidence intervals within brackets

Whereas the weight for EQ increases to 0.34 for the scenarios with the low HH reference, and the weight for NRandES increases to 0.27.

Based on a Mann-Whitney *U* test, the weight of the HH AoP is higher for the respondents with choice cards using the high HH reference ($p = 0.000$), which corroborates the conclusion of the econometric study.

More insights can be derived when looking at the results within each income group for different HH references (*italics font* in Table 11). For the high-income group and high HH reference, HH has the highest weight (0.39), followed by EQ (0.36), and then by NRandES (0.25). When using the low HH reference, EQ has the highest weight (0.41), followed by HH (0.32), and then NRandES (0.27). It is thus apparent that the use of the different HH references has a substantial influence on the prioritization of the responses in high-income countries.

The overall weight for the HH AoP in the upper- and lower-middle-income groups is the same (0.39). However, the HH references have a different effect on the obtained weights. In the upper-middle-income group, there is a difference of 0.03 between the weight for HH with high and low HH references. In the lower-middle-income group, there is no difference, showing the lowest sensitivity for the use of the different HH references in the whole set of responses. Table 11 shows that only the lower-middle-income group with the low HH reference has a higher weight for NRandES (0.31) than EQ (0.30).

Piece-wise linear value function model The results from the nonlinear (piece-wise) value function model are very similar to the linear case, and so they are only briefly summarized here (further results are available in Annex I and J in the ESM). The mean weights for this model grouped by HH reference are presented in Table 12. As for the linear model, the weight of the HH AoP is higher for the group where the health reference is high ($p = 0.000$).

The same order of priorities for the AoPs is preserved with the high and low HH references, as well as when they are combined. When combined, the HH AoP is weighted

Table 12 Mean weights grouped by HH reference for all income groups (MCDA piece-wise linear model)

MCDA piece-wise linear	**Human health**			**Ecosystem quality**			**Natural resources and ecosystem services**		
	$s_1^{<ref}$	$s_1^{>ref}$	w_1	$s_2^{<ref}$	$s_2^{>ref}$	w_2	$s_3^{<ref}$	$s_3^{>ref}$	w_3
HH reference = high	0.21	0.21	0.43 [0.41, 0.44]	0.11	0.19	0.30 [0.29, 0.31]	0.11	0.17	0.27 [0.26, 0.29]
HH reference = low	0.19	0.2	0.39 [0.37, 0.40]	0.13	0.2	0.34 [0.32, 0.35]	0.11	0.17	0.28 [0.27, 0.29]

Ninety-five percent (95%) confidence intervals within brackets

$s_k^{<ref}$ slope below reference; $s_k^{>ref}$ slope above reference; w_k implicit weight in the range and 95% confidence intervals within brackets

Table 13 Mean weights clustered by income groups (MCDA piece-wise linear model)

MCDA piece-wise linear	Human health			Ecosystem quality			Natural resources and ecosystem services		
	$s_1^{<ref}$	$s_1^{>ref}$	w_1	$s_2^{<ref}$	$s_2^{>ref}$	w_2	$s_3^{<ref}$	$s_3^{>ref}$	w_3
All income groups	0.20	0.21	0.41 [0.40, 0.41]	0.12	0.20	0.32 [0.31, 0.33]	0.11	0.17	0.28 [0.27, 0.29]
High-income group	0.16	0.21	0.36 [0.34, 0.38]	0.15	0.23	0.38 [0.36, 0.40]	0.11	0.15	0.26 [0.24, 0.27]
Upper-middle income group	0.18	0.21	0.39 [0.37, 0.40]	0.12	0.21	0.33 [0.31, 0.34]	0.1	0.18	0.29 [0.27, 0.30]
Lower-middle income group	0.15	0.22	0.37 [0.35, 0.40]	0.08	0.21	0.30 [0.28, 0.32]	0.09	0.24	0.33 [0.31, 0.35]
Low-income group	0.29	0.19	0.48 [0.46, 0.49]	0.12	0.15	0.27 [0.26, 0.28]	0.12	0.13	0.25 [0.24, 0.27]

Ninety-five percent (95%) confidence intervals within brackets

$s_k^{<ref}$ slope below reference; $s_k^{>ref}$ slope above reference; w_k implicit weight in the range and 95% confidence intervals within brackets

highest with a weight of 0.406, followed by the EQ AoP with a weight of 0.318, and the NRandES AoP with a weight of 0.277. When comparing the weights from the linear to the piece-wise linear models, the differences between the values for the weights when the high or low HH references are used are slightly reduced for HH (by 0.004) and NRandES (by 0.006), but slightly increased (by 0.002) for NRandES.

Further considerations can be derived from comparing the weights grouped by income and HH reference, as shown in Table 11. The weights obtained by combining the results from the scenarios with both HH references (bold font in Table 11) are 0.41 for HH, 0.32 for EQ, and 0.27 for NRandES. Differences between these weights and the weights obtained from the econometric model are small (approximately 0.01).

In the high-income group, the EQ AoP has the highest weight (0.39). This is followed by the HH AoP (0.36) and lastly by the NRandES AoP (0.26). Again, the high-income group is the only one that shows this preference order for the AoPs, as the HH AoP receives the highest weight for the other income groups. The weights ranking found for the lower-middle income group (highest weight for the HH AoP, 0.39, followed by weights that are 0.31 for the other two AoPs) and for the low-income group (highest weight for the HH AoP, 0.48, which is much higher than the one for the EQ AoP and the NRandES AoP, i.e. at least a 0.2 gap between HH and the next highest weighted AoP) also match the ranking of the results from the econometric model. The two models differ on the upper-middle income group, with the additive value function model giving a higher weight for the HH AoP, followed by EQ and NRandES. Based on the Kruskal-Wallis H test, all the weights are different among the income level groups, when considered simultaneously ($p=0.000$).

The weights grouped by income level are found in Table 13. Due to the similarity with the findings from the linear model, only the weights obtained by combining the results from the scenarios with both HH references are shown. It is confirmed that the highest weight for the high-income group is for EQ, with a weight of 0.38. The HH and NRandES AoPs follow sequentially, with weights of 0.36 and 0.26, respectively. As in the linear model, the HH AoP has the highest weight in all the other income groups, and the EQ and NRandES AoPs are in the second and third position in the upper-middle and lower-middle-income groups, respectively. The only difference from the linear model is that, in the lower-middle-income group, the EQ and NRandES AoPs do not have the same weight. Instead, in the piece-wise linear model the NRandES AoP has a higher weight (0.331) than the EQ (0.296).

The piece-wise linear model results show that for the EQ and NRandES AoPs, a difference of impacts in the range better than the reference has less relevance (importance) than a similar difference of impacts in the range worse than the reference ($p=0.000$). The same is not true for the HH AoP ($p=0.140$). This is represented in Fig. 6, where the value function for the HH AoP is close to linear over the whole spectrum of the impact.

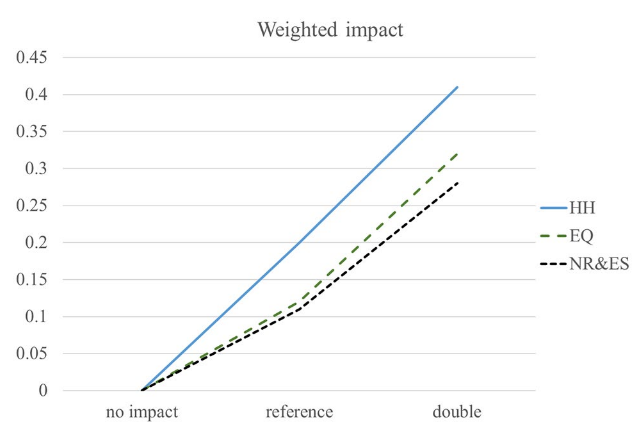

Fig. 6 Mean value functions for the piece-wise linear model for each AoP

Table 14 Proportion of the world's population living in countries with different income levels and % of valid responses used in the econometric and the MCDA approaches

Income group	%	% of valid responses for the econometric approach	% of valid responses for the MCDA approach
High-income	15.78	24.85	25.31
Upper-middle-income	31.85	29.71	27.64
Lower-middle	43.23	16.04	16.78
Low	9.14	29.39	30.26

3.1.3 World population share-adjusted global weights

The global weights calculated with the econometric and MCDA disaggregation approaches are based on all the survey responses. However, as summarized in Table 14, the proportions of the world's population per income group, according to the World Bank population data from 2021 (World Bank 2023e), show a distribution that does not match the proportion of survey respondents from each country income level group. For this reason, weights that account for the share of the world population per income group are calculated as well.

The weights obtained for each of the income groups from Table 9 (econometric) and Table 11 (MCDA) are the basis for the calculations. These weights are combined with the income level shares of the world's population (from Table 14 above), so that population-adjusted weights are calculated (see Table 15). An example of how this calculation is performed for the HH econometric population adjusted global weight is as follows: $HH_{econometric} = (Population_{high\text{-}income} \times HH_{weight\ high\text{-}income}) + (Population_{upper\text{-}middle\text{-}income} \times HH_{weight\ upper\text{-}middle\text{-}income}) + (Population_{lower\text{-}middle\text{-}income} \times HH_{weight\ lower\text{-}middle\text{-}income}) + (Population_{low\text{-}income} \times HH_{weight\ low\text{-}income}) = (0.1578 \times 0.34) + (0.3185 \times 0.36) + (0.4323 \times 0.36) + (0.0914 \times 0.54) = 0.37$.

These weights can be a solution to account for the fact that the sample sizes from the different countries do not accurately reflect the global population/income level distribution. If the opinions expressed by respondents are on average representative for the world's population of their given income level, then the weights obtained using this adjustment would be more representative for the world's population than the weights provided in Tables 8, 9, 10, 11, 12, and 13. Table 9 shows the econometric model results for all income groups and both human health references, which provide the following weights for HH, EQ, and NRandES, respectively: 0.42, 0.31, and 0.26. The corresponding weights from the MCDA linear model are 0.41, 0.32, and 0.27. Comparison with the population share-adjusted results in Table 15 shows that these adjusted results give generally lower weights for HH, higher weights for EQ, and higher weights for NRandES. This is due to the differences in proportions of the global population compared to the respondent population answering the survey. The global population has a relatively low proportion of low-income population (9.14%), compared to the low-income proportion of the respondents to the survey (29% for econometric and 30% for MCDA[4]). The number of lower-middle income responses in the survey was 16–17%, whereas they account for 43% of the global population. For upper-middle income, the proportion of global population is 32%, which is not too far from our survey population of 28–30%. The high-income respondents represented 25% of the respondent population, while only 16% of the global population is in high-income countries.

4 Discussion

4.1 Applicability and timeframe of validity

The weights presented in this paper are suitable for those projects with the same AoPs as GLAM and when no weights have been elicited by the analysts. This absence of elicited weights can occur mainly for the following reasons:

Table 15 Population share-adjusted global weights

AoP	Econometric (HH reference high + low)	MCDA linear (HH reference high + low)
HH	0.37	0.39
EQ	0.34	0.33
NRandES	0.29	0.29

[4] The econometric approach included 3198 respondents that had answered at least one choice task, so the proportions are slightly different to the MCDA disaggregation approach that only analysed responses from respondents that completed all nine choice tasks, i.e. 3003.

1. The analysis is being conducted for decision-making purposes, e.g. a company deciding between different ways to manufacture a product, and the decision-makers would rather not use their personal preferences. The decision-makers might be unavailable or feel incapable of providing such weights, or they might wish to show that conclusions are independent from their own preferences. Therefore, the analysts would rather use external weights, in this case, weights that represent the preferences of the respondents of a large global survey.
2. The analysis is not being conducted for decision-making purposes, and no decision-makers are involved in the analysis. For instance, the analysts are assessing some product(s) or system(s) for a research paper. The analysts have computed the environmental impacts for the AoPs, and they would like to synthesize them as a way of offering comparable AoPs values and/or aggregated results for the readers of their work.
3. In situations 1 or 2 above, the analysts consider engagement with stakeholders or a specific population to elicit weights, but resources (time, money) are insufficient to provide a specific weighting set for the project.

In these cases, common weighting choices are the use of equal weights or some default weights, if these are available. *The use of these weights is ideally suited for these types of decision-making problems, namely where the decision makers do not want to or have no resources to identify a set of weights themselves, or when decision makers are not involved.*

There is no pre-established timeframe of validity for these weights. It is common practice in LCA that temporal characteristics are analysed by identifying technology shift (Henriksen et al 2020), which means that datasets are valid as the best available data until an update is proposed based on new research. The authors suggest this approach for these factors. However, it should be noted that these weights are not life cycle inventory data, associated with emissions from technological processes (as the data described in Henriksen et al. (2020)). Another temporal aspect is that impacts calculated in LCA studies are often calculated with a time perspective included in the impact calculations. Thus, impacts calculated over the whole timeframe (e.g. often 100 years for global warming impacts) will not have a uniform effect over that timescale. Although the potential impact results at endpoint for a given LCA study are presented as DALY, PDF_{GLO}.year, and a USD value with a reference year, they do not necessarily happen in a specific year, but rather as a result of the system in focus for the analysis. Temporal issues related to how impacts are calculated and the temporal scale are part of the goal and scope of the given LCA study.

It is also important to note that these weights, as any other set of weights, cannot be used in situations where there is a need to perform and publish an ISO-compliant study, as according to the ISO standard on LCA (ISO 14044 2006), it is not possible to compare product systems and disclose the weighted single score results to the public.

Table 16 Conceptualization of LCA, environmental science, and economics perspectives in relation to weighting and aggregation in operations research, including the weights that are presented in this paper

LCAs	Environmental science/ economics	Weights type	Aggregation	Value functions	Weights presented in this paper
Individualist	Weak sustainability	Compensation rates	Additive (e.g. weighted arithmetic mean)	Linear	Yes: econometric, MCDA linear value functions
Hierarchist	Between weak and strong sustainability	Compensation rates	Additive (e.g. weighted arithmetic mean)	Nonlinear	Yes*: MCDA piece-wise linear value functions penalizing losses over gains
		Compensation rates	Non-additive (e.g. multilinear and multiplicative models, weighted geometric/harmonic mean, Choquet integral, ordered weighted average)	Linear or nonlinear	No
		Importance coefficients	Outranking (e.g. ELECTRE, PROMETHEE)	NA	No
Egalitarian	Strong (critical) sustainability	No weighting	Aggregated impact = worst impact (MIN operator)	NA	NA

*Achievable also with the econometric approach, but not a focus of this research

Table 17 Hypothetical example of the application of the weights from the econometric approach

	HH	EQ	NRandES	Total
Impacts (Inv × CF)	8.9 E^{-5}	3.9 E^{-13}	1.56 E^{-4}	
Impacts (units)	DALY/FU	PDFglo.yr/FU	USD/FU	
NV[a]	0.101	1.57 E^{-11}	841	
NV (units)	DALY/person.yr	PDFglo.yr/person.yr	USD/person.yr	
Normalized impact score: impacts/NVs	8.81 E^{-4}	2.49 E^{-2}	1.85 E^{-7}	
Normalized impact score: impacts/NVs (units)	Person.yr/FU	Person.yr/FU	Person.yr/FU	
Weights	0.42	0.31	0.26	
Single score with non-dimensional weights	0.00037010	0.00771924	0.00000005	0.008
Single score with non-dimensional weights (units)	Person.yr/FU	Person.yr/FU	Person.yr/FU	Person.yr/FU
Impact contribution to single score (%)	4.575%	95.424%	0.001%	100%

[a]0.101 is equal to 37 days/365 days; 1.57 E^{-11} is calculated as 0.12 divided by world population; 841 is the 7.5% of the average per capita GDP for 2018

4.2 Where do the weights fit in the LCA domain and beyond?

The weighting subtask also developed a conceptualization of how different methodological approaches in LCA, environmental science, and economics can be related to weighting and aggregation in operations research. This is presented in Table 16, together with what has been achieved via the work presented in this paper.

The cultural perspectives of individualist, hierarchist, and egalitarian in LCA (Hofstetter 2000) are related to the conceptualizations of sustainability going from weak to strong (Victor et al. 1998). From an operations research perspective, the use of an individualist (weak sustainability) and hierarchist (between weak and strong sustainability) perspective implies compensation between the AoP, which fits well with weights used as compensation rates and an additive aggregation model (Munda 2016). Decision analysts interested in exploring less compensatory perspectives, meaning stronger sustainability frameworks, should instead transition to non-additive aggregation models or could employ different aggregation models such as the outranking methods (Cinelli et al. 2014; Dias 2021).

4.3 Application of the weights in LCA practice

The weights proposed by this subtask for both the econometric and the MCDA disaggregation (linear) approaches are adimensional, between 0 and 1, and they can be used by LCA practitioners (e.g. consultants) and researchers to calculate single scores for environmental impacts of products and processes, using Eq. 7:

$$I = \sum_{\text{Endpoint}} \sum_{\text{Impact}} \sum_{x} \left(\frac{\text{Inv}(x) \times \text{CF}(\text{endpoint}, x)}{\text{NV}(\text{endpoint})} \times \text{WF}(\text{endpoint}) \right) \quad (7)$$

where:

I is the result of the overall environmental impact
Inv(x) is inventory data for each input x used for the target of the assessment
CF(endpoint, x) is the characterization factor for each input x used for the target of the assessment
NV(endpoint) is the normalization value for the given endpoint
WF(endpoint) is the weighting factor (= weight) for the given endpoint

Equation 7 belongs to the endpoint type 1 of weighting as presented in a chapter dedicated to weighting in LCA (Itsubo 2015), and it fits with the individualist LCA perspective presented in Table 16 in the previous section. The weights proposed for this endpoint type 1 of weighting can be used for two purposes, (i) to assess a single product (or system) to identify hotspots in life cycle stages and (ii) to compare different products (or variants for a system) to identify which one performs better. This proposed endpoint type of weighting is independent from the set of alternatives (products, services) that are being evaluated. This means that in cases where alternatives are added or deleted from the set, the overall environmental impact will not change. Consequently, this endpoint type of weighting is not affected by rank reversal.

An example for the application of the weights from the econometric approach for a case study with a single alternative is presented in Table 17. The first stage is the identification of the impacts caused by the alternative under study. All of the new GLAM LCIA methods have not yet been used in any published case studies, so a case study has been developed for tinned tuna in brine, based on results in Helias et al. (2023) and De Vlieghere et al. (2023). This has been done to show an example of the LCIA results that could be expected in a real-life case study.[5] The results taken from the case study in Helias et al. (2023) are those for yellowfin tuna (i.e. global impacts on biodiversity 3.9 E^{-10} PDF_{GLO}.year per tonne). The results for tuna in brine from Fig. 4 in De Vlieghere et al. (2023) were used for the other impact categories. As they are presented as midpoint indicators in De Vlieghere et al. (2023), a calculation was performed to provide endpoints similar to those expected using the GLAM AoP methodology from the midpoint indicators. The AoP results are presented per functional unit (FU, 1 kg of tinned yellowfin tuna in brine), i.e. impact values of 8.9 E^{-5} DALY/FU for HH AoP, 3.9 E^{-13} PDFglo.year/FU for EQ AoP and 1.56 E^{-4} USD/FU for NRandES AoP.

The second stage involves the selection of the NVs for the AoPs, which in this case are the ones provided by the normalization subtask of the GLAM project, all expressed in relation to a "person.year". For HH AoP, this value is equal to 0.101 DALY/person.year, and it is the ratio between 37 days lost per person (i.e. average between 19 and 55 days, low and high HH reference, respectively) and the days in a year (i.e. 365). For EQ AoP, the value is 1.57 E^{-11} PDFglo.year/person.year, and it is calculated as 0.12 PDFglo.year divided by the world population. Lastly, for the NRandES, the NV is 841 USD/person.year, and it is 7.5% of the average per capita GDP for 2018 (the year used by the NRandES subtask for the calculation of the raw NV) per year.

The normalized impact score (i.e. impacts/NVs) is then obtained for each AoP by dividing the impacts with the respective NV. For example, the normalized impact score of 8.81 E^{-4} person.year/FU for HH AoP is calculated as:

$$\frac{8.9E^{-5}\ \mathrm{DALY/FU}}{0.101\ \mathrm{DALY/person.year}} = 8.81\ E^{-4}\ \mathrm{person.year/FU}$$

The weights used for this example are the ones obtained from the econometric approach by accounting for the preferences from all the income groups (from Table 9), resulting in 0.42 for HH AoP, 0.31 for EQ AoP, and 0.26 for NRandES AoP.

The normalized impact scores are then multiplied by the weights, resulting in the individual contributions of the single score to the overall environmental impact of the alternative under assessment.

[5] The chosen case study is a simplified version of the cited work by Helias et al. (2023) and De Vlieghere et al. (2023), and it should not be used for any other purposes than the example of application of the weighting approach in this specific paper.

For example, the single score of 0.00037010 person.year/FU for HH AoP is calculated as:

$$8.81\ E^{-4}\ \mathrm{person.year/FU} \times 0.42 = 0.00037010\ \mathrm{person.year/FU}$$

The individual contributions from the single score per AoP can then be summed up to provide the overall impact of the alternative under evaluation, which in this case is calculated as:

$$0.00037010\ \mathrm{person.year/FU} + 0.00771924\ \mathrm{person.year/FU}$$
$$+0.00000005\ \mathrm{person.year/FU} = 0.008\ \mathrm{person.year/FU}$$

For this case study, it is evident how the largest share (over 95%) of the impacts is determined by the EQ AoP, followed by the HH AoP at just over 4.5%, and with a negligible contribution close to 0% for the NRandES AoP.

4.4 Comparison of the new set of weights with existing weighting sets

The weighting approaches applied in the GLAM project, LIME3 (Itsubo et al. 2018), and Ecoindicator99 (EI99) (Goedkoop and Spriensma 2001) are all endpoint-based and can be compared to each other to a certain extent. As shown in Table 18, this study (GLAM) and EI99 have three comparable AoPs, while LIME3 has four AoPs including social assets as the fourth AoP. The natural resources AoP in LIME3 is linked to fossil fuels and does not include ecosystem services. Social assets in LIME3 are described as "valuables in human society such as fishery, agriculture, and forestry". Consequently, the combined LIME3 AoPs of natural resources and social assets can be considered similar to the GLAM NRandES AoP. The resources AoP in EI99 is limited to minerals and fossil fuels only.

The first AoP is the same in all three methods: i.e. "Human Health", measured in disability-adjusted life years (DALY) and considering both morbidity and mortality. The name of the second AoP is the same in both GLAM and EI99: "Ecosystem Quality", expressed as % of species, while in the LIME3 method, it is expressed as "Biodiversity" using the unit of "expected increase in number of extinct species (EINES)". Although the EQ AoP has the same name for GLAM and EI99 impact assessment methods, GLAM defines this AoP as "percentage of terrestrial species (mammals, amphibians, birds, reptiles, and vascular plants) that are put at risk of extinction", while EI99 describes EQ as "percentage of vascular plant species that have disappeared in a certain area during a certain time". Thus, in the GLAM project, a wider range of terrestrial species is used to calculate the NV.

The third AoP is named differently in all three methods: "Natural Resources and Ecosystem Services" (NRandES)

Table 18 Comparison of the new weights with existing weighting sets

	This study	LIME3	EI99
AoP 1	**Human health**	**Human health**	**Human health**
Unit	DALY	DALY	DALY
Expression	Days of healthy life lost per person per year	Loss of life expectancy	Number of life year lost and lived disabled
AoP 2	**Ecosystem quality**	**Biodiversity**	**Ecosystem quality**
Unit	%	EINES	% vasc. plant species $*km^2$ *yr
Expression	% of global terrestrial species that are put at risk of extinction	Loss of 1 specie	% of species that have disappeared
AoP 3	**Natural resources and ecosystem services**	**Social assets**	**Resources**
Unit	USD	USD or JPY	MJ
Expression	USD loss of natural resources and services from the ecosystems per person per year	Loss of natural resources per person	Surplus energy needed for future extraction of minerals and fossil fuels
AoP 4		**Primary production**	
Unit		Billion tons	
Expression		Loss of forest	
Survey group	4 income groups	G20 countries	1 panel
Sampling	Purposive + random + snowball	Random + quota	Purposive
Number of respondents	From 513 to 950 per income group	200–250 for emerging, 500–600 for developed	82
Method	DCE and MCDA disaggregation based on DCE	Conjoint Analysis based on DCE	Ranking

in GLAM, "Social Assets" in LIME3, and "Resources" in EI99. GLAM and LIME3 use the unit of USD for this AoP, while EI99 uses MJ surplus energy. In GLAM, NRandES AoP includes discovered minerals, fossil fuels, as well as ecosystem services that support everyday life, including plant growth, water, and food sources.

LIME3 is the only method to have a fourth AoP, i.e. "Primary Production", defined as loss of forest expressed in the unit of billion tons of forest per year.

In terms of survey design, EI99 used an expert panel weighting method, while LIME3 and GLAM used DCE to provide data for the weight calculations. EI99 asked 365

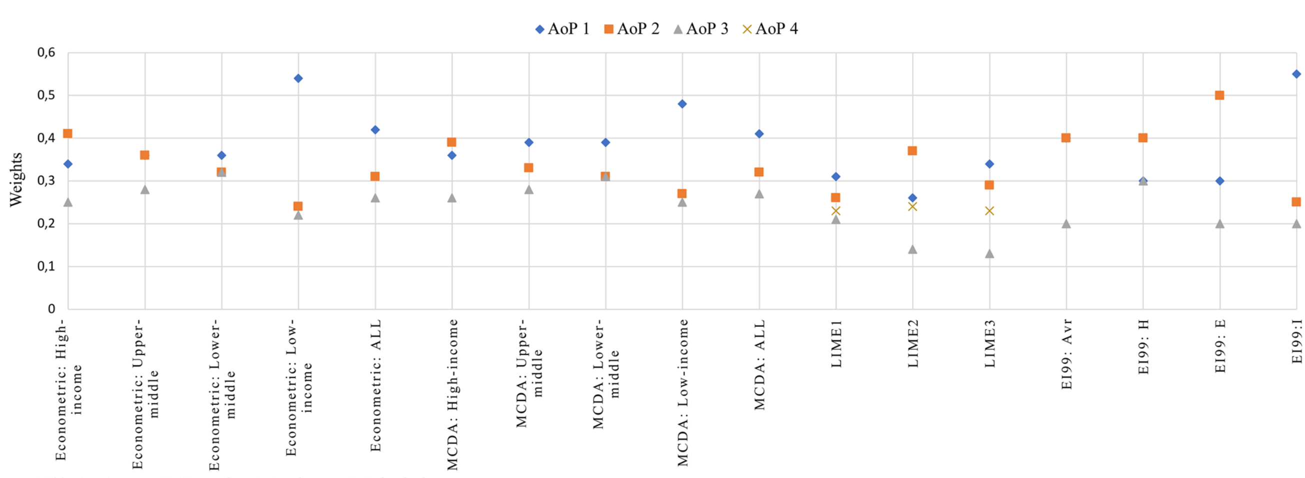

Fig. 7 Comparison of the new weights and existing weighting sets

LCA experts who had attended a Swiss discussion forum (obtained 82 responses) to rank the AoPs, while GLAM and LIME3 asked citizens to choose one scenario out of three using a DCE format.

Figure 7 reports the weights calculated in GLAM, LIME1 (Itsubo et al. 2004), LIME2 (Itsubo et al. 2012), LIME3 (Murakami et al. 2018), and EI99 (Goedkoop and Spriensma 2001). EI99 is the oldest method and uses cultural theory and divides the weighting into three groups: individualist (I), egalitarian (E), and hierarchist (H). It also gives an average value for these three groups. Individualists prioritize human health ($WF_{HH}=0.55$), while the egalitarians give more importance to EQ ($WF_{EQ}=0.50$). As an average, EI99 weighted HH 0.40, EQ 0.40, and resources 0.20. The high-income country respondents to the GLAM DCE can thus be described as displaying egalitarian traits, whereas lower income country groups are more individualist.

Weights in LIME1, LIME2, and LIME3 are 0.31, 0.26, and 0.34 for HH, 0.26, 0.37, and 0.29 for biodiversity, 0.23, 0.24, and 0.23 for primary production, and 0.21, 0.14, and 0.13 for social assets, respectively. Weights for HH and biodiversity increased from LIME1 to LIME3 at the expense of social assets, while primary production remained the same.

In GLAM, there are two approaches, econometric and MCDA disaggregation. Weights are calculated for low NVs (19 days) and high NVs (55 days) for HH. Also, a combined set of weights is calculated for both approaches.

For the econometric (MCDA linear) approach, HH weights for HH_{Low} is 0.40 (0.39), HH_{High} is 0.45 (0.43), and HH_{All} is 0.42 (0.41). EQ weights in the econometric (MCDA linear) approach are 0.33 (0.34) for HH_{Low}, 0.30 (0.31) for HH_{High}, and 0.31 (0.32) for HH_{All}. For NRandES, the weights are 0.28 (0.27), 0.25 (0.26), and 0.26 (0.27) for HH_{Low}, HH_{High}, and HH_{All} in econometric (MCDA linear) approach.

As seen in Fig. 7, weights of combined data for both approaches, econometric, and MCDA, (0.42 and 0.41 for HH, 0.31 and 0.32 for EQ, 0.26 and 0.27 for NRandES, respectively) are similar for GLAM. However, there is a 0.11 difference in HH between the econometric approach and LIME1, 0.08 differences with LIME3, while this difference is only 0.02 between the GLAM econometric approach and EI99's average weighting for HH.

For EQ, weights calculated as part of the GLAM project are between the LIME3 and the EI99 (average) weight values. The weights for EQ for the high HH normalization values (0.30 for both approaches) in the GLAM project are very close to the LIME3 weights (0.29) for EQ (the difference is only 0.01).

For the NRandES (or resources in EI99 or social assets in LIME), the weights from the GLAM project are the highest among the methods compared. However, as GLAM NRandES represents what LIME defines as social assets and natural resources, they could be combined to be a comparable AoP to the NRandES for GLAM and resources for EI99. This would make the total weighting for LIME1 0.44, LIME2 0.38, and LIME3 0.36. The combined weighting of LIME1, LIME2, and LIME3 would be higher than GLAM NRandES (ranging from 0.25 to 0.28) and higher than the EI99 average weighting for resources (0.20). However, this requires acknowledging the likelihood of a splitting bias occurring (Marttunen et al. 2018) (i.e. dividing an objective into two sub-objectives leads to an increase in the weight of the sum of the two). This means that it is likely that had the NRandES AoP in the GLAM DCE been split into two AoPs (as for LIME), then they would possibly have been assigned a higher combined weight.

This study is the only one that includes respondents from low-income countries. For the HH AoP, the weights elicited in the LIME studies are more similar to those for the three higher income groups than in this study. This is not surprising, as LIME surveyed respondents in G20 countries.[6]

Lastly, the absolute differences between the weights in the econometric approach, when considering HH_{Low} and HH_{High}, is 0.11 between HH and EQ, and it is 0.5 between EQ and NRandES. A similar trend is visible for the MCDA approach. In addition, the confidence intervals in Tables 9 (econometric) and 11 (MCDA) indicate that the weights are different from equal weighting (1/3, 1/3, 1/3) in a statistically significant way. A case study with weights that differ notably from equal values is one that deals with an LCA where the weighting of focus is for a low-income group country. In this case, the weights from the econometric approach for the HH are between 0.49 (low HH reference) and 0.58 (high HH reference), while the ones for the remaining AoPs are much lower, reaching at most 0.26. A similar trend is confirmed for this type of case study for the weights from the MCDA approach.

4.5 Limitations of the study and recommendations for future research

There are six main limitations that must be acknowledged as part of this research. *First*, the sample of respondents included in the survey is not fully representative for the population of each income group. This is due to the number of responses received, as well as the combination of different sampling methods. Figure 8 shows the responses received

[6] The G20 countries are Argentina (upper-middle income), Australia (high), Brazil (upper-middle), Canada (high), China (upper-middle), France (high), Germany (high), India (lower-middle), Indonesia (lower-middle), (high), South Africa (upper-middle), Türkiye (upper-middle), UK (high), USA (high), and European Union.

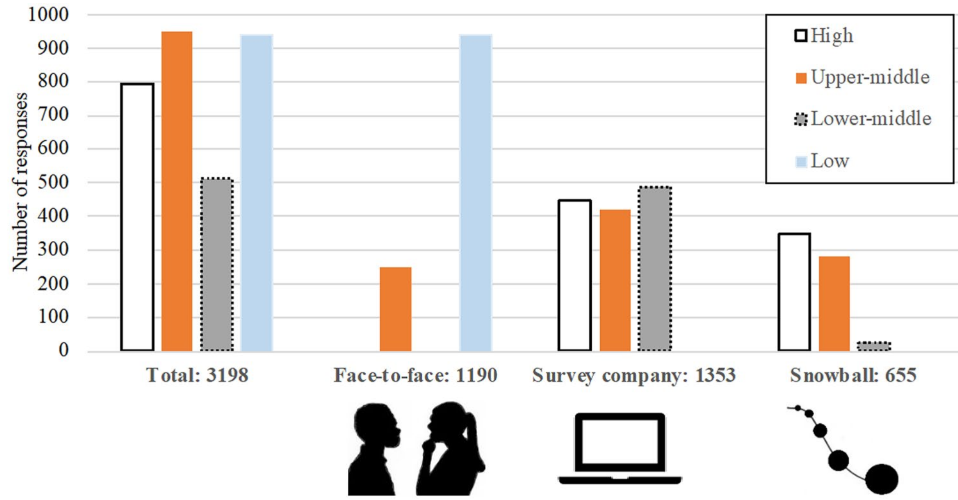

Fig. 8 Number of responses for the different sampling strategies and income groups

with the different sampling strategies. Different levels of sampling rigour are apparent, with high quality of sampling achieved with face-to-face interviews in Uganda, Burkina Faso, and partly in Türkiye (36.8% of total responses). Medium quality sampling was achieved with targeted respondents from a survey company list for Japan (Tokyo, high-income), China (Shanghai, upper-middle-income), and India (Mumbai, lower-middle-income) (42% of responses), and low sampling quality was reached with the remaining participants via the snowball method (21.2% of responses).

Second, the weights proposed in this paper are not fit for LCAs that are focused on the assessment of the impacts at a local or regional scale. In such cases, weighting sets based on that scale, from the affected population and with impacts expressed as AoPs that are geographically disaggregated, would be preferred.

Third, the models used for the calculation of the weights assume two types of value functions as representative for the preferences of the respondents. These are linear (both econometric and MCDA disaggregation approach) and piece-wise linear (only the MCDA disaggregation approach). Weights for other value function shapes have not been estimated.

Fourth, the models used for the calculation of the weights assume that an additive aggregation is applied to derive the overall environmental impact. Weights obtained from assuming non-additive aggregation models (e.g. multilinear and multiplicative models, weighted geometric/harmonic mean, Choquet integral, ordered weighted average) have not been tested (see Section 4.2 for more details about different aggregation algorithms).

Fifth, the weights that have been obtained are likely to be dependent on the NVs used in the survey. Consequently, different levels of recommendations apply according to the NVs that they will be used with. A high level of recommendation applies to the weights when they are used with the NVs they have been developed with. A medium level of recommendation applies to the weights if they are used with NVs that are within the upper boundaries of the NVs tested in the GLAM survey. This means 110 days per person per year for HH, 24% of species at risk of extinction, and 6796 USD lost per person per year (for high-income). A low level of recommendation is applicable if the weights are used with NVs that are outside the upper boundaries of the NVs tested in the GLAM survey.

Sixth, separating AoPs into three separate categories of HH, EQ, and NRandES can imply that human beings are separate from nature and risk underplaying the importance of the many interactions between social and natural systems. A DCE, such as the one presented in this study, where people are required to make trade-offs between these different AoPs can neglect indirect and long-term effects, like future generations being affected by ecosystem damage (i.e. reductions in ecosystem quality and ecosystem services).

Future research to extend the work presented in this paper could include the use of nonlinear value functions as well as other (e.g. non-additive) aggregation models (Langhans et al. 2014). Larger sample sizes of respondents, targeting more countries in each income category, in a randomized, representative fashion would also improve the robustness of the response data used to calculate the weights.

A new survey could also be designed to clearly lead the respondents to reason in relative terms (with respect to the NVs), so that weights would need no adjustment if the NVs change, and this change is not too large. The population-adjusted weights presented in Section 3.1.3 are one way of adjusting the responses to better reflect the world's population. A revised design of the survey, requiring more resources, would also involve testing different survey layouts, graphics to represent the impacts and wording options, e.g. focus groups using texts in the local languages, and different wordings of the impacts in the AoPs. Monetary values could be introduced by including a willingness to pay approach, or an environmental tax (as in the LIME approach) (Murakami et al. 2018).

Another avenue for future research could include testing DCE and MCDA (disaggregation) to elicit weights for LCA midpoint indicators instead of endpoint ones, as in the GLAM work. It must however be noted that the survey design and task of the respondents would be much more complex as there would be a large number of attributes needed to describe the scenarios (e.g. more than 10 for the EU product environmental footprint (PEF)) (Sala and Cerutti 2018). Furthermore, the communication of the meaning of the midpoint impacts to “regular” citizens would be a major challenge, due to technical terminology for the units of impact (e.g. kg CFC-11 eq., CTUh, kBq U-235 eq.). A promising opportunity for future (survey) research on weighting in LCA would be the inclusion of the interdependencies among the AoPs, especially when accounting for different time horizons to characterize the impacts (Lueddeckens et al. 2020).

As far as robustness is concerned, in addition to the econometric and MCDA methods used in this study, other weighting methods suitable to the GLAM project (lower right of Fig. 1) could be applied to calculate the weights. As mentioned in Section 2.1, there were also members of the weighting subtask who had reservations on the chosen methods to calculate the weights. They agreed to transfer their reservations in the form of recommendations for future research. These included calculating the weights according to (i) planetary boundaries (Richardson et al. 2023), (ii) the needs of future generations, and (iii) global burden of disease (IHME 2019). Furthermore, a combination of distance to target and monetary methods approach to weighting was also proposed during the subtask work in this phase of the GLAM project, hereafter called the combined method. The aim of this was to incorporate a strong sustainability perspective, where the first condition of satisfying the needs of the present generation is quantified via willingness to pay (because we can measure it and it is exchangeable), and the second condition is the cost for “not compromising” the ability of future generations to satisfy their needs, i.e. costs for the present generation for the preservation of natural resources.

5 Conclusions

This paper has presented an approach for eliciting population’s preferences in order to calculate weights for use in the optional weighting step in LCA. The work was performed as part of the activity of the weighting subtask of the normalization, weighting, and cross-cutting issues task force in the third phase of the United Nations Environment Programme (UNEP) Life Cycle Initiative’s “Global Guidance for Life Cycle Impact Assessment Indicators and Methods” (GLAM) project.

Weighting is an optional part of LCIA methodology, but useful for situations where the LCIA provides decision-makers with results that do not all point in the same direction. In such cases, the decision-maker is faced with a choice about which AoPs they value more in endpoint-based LCA, or which impact categories they would like to prioritize in midpoint-based LCA. Without the availability of weighting factors, decision-makers can consciously, or sub-consciously weigh each AoP (or whichever impact categories they use) equally. They can also be subject to criticism about choosing weighting that favours their pre-defined preferences. The weights presented in this paper are focused on the AoPs of the GLAM methodology, and they are suitable for decision-making situations where the problem owner does not want to or cannot identify a set of weights themselves, or when a problem owner does not exist (typically in the context of academic projects).

This study provides sets of weights for the three AoPs (human health (HH), ecosystem quality (EQ), natural resources and ecosystem services (NRandES)), of the GLAM methodology based on a discrete choice experiment that gathered data about the preferences of a large number (3198) of global respondents, the largest so far in any such study. The response data was obtained from a subset of countries from each income level defined by the World Bank (i.e. low, lower-middle, upper-middle, and high). There was reasonably extensive sampling from six countries, as well as more limited numbers of responses from a range of other countries across the income groups. This is the first time a weighting study of this kind has included such an extensive sample of low-income country opinions. The most extensive peer-reviewed similar studies are the two LIME surveys that have achieved similar sampling rates per country, but only within countries that are part of the G20.

The adimensional weights presented in this paper were calculated using two different approaches: econometric and MCDA. Since the survey was designed to fit the econometric approach, the authors suggest the use of these factors (provided in Table 9) if the user wishes to choose between these. The use of the two calculation approaches demonstrates the robustness of the weights derived. This is confirmed by the small differences in the weights derived for all income groups using both weighting approaches (i.e. HH 0.42, EQ 0.31, and NRandES 0.26 for the econometric model in Table 9 compared to HH 0.41, EQ 0.32, and NRandES 0.27 for the MCDA linear model in Table 11). It is possible to observe some differences in preferences across the income groups. For both weighting approaches, the high-income group weights the EQ AoP higher than HH, whereas for low-income countries, HH is assigned a higher weight than the other AoPs. The NRandES AoP is consistently weighted lower than the other AoPs, independent of income level, with the exception of lower-middle-income countries (where EQ has a lower weight than NRandES using the econometric approach and the MCDA disaggregation approach when the HH reference is low).

The sets of weights presented in this paper are also shown in relation to existing LCA, environmental science, and economics perspectives. The econometric and MCDA weights derived using linear functions fall under the individualist (LCA) and weak sustainability (environmental science /economics) perspectives. These weights are compensation rates; their aggregation is additive, and they are based on linear value functions. The MCDA weights derived using piecewise linear value functions can be considered to fall under the hierarchist (LCA) and between weak and strong sustainability (environmental science/economics) perspectives. These weights are also compensation rates, with additive aggregation, but they are based on nonlinear value functions that penalize impacts above the NV. This study shows a key contribution of econometric and MCDA to the practice of LCA, especially by enabling decision-makers who do not want to or cannot identify a set of weights themselves to advance the interpretation of their LCA results.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11367-024-02330-w.

Acknowledgements The authors thank all the members of the Weighting subtask for the constructive input provided during the development of the survey, as well as the (co)-chairs of the GLAM task forces for coordinating the preparation of the normalization values used to develop the reference scenarios.

Author contribution Conceptualization: Ayşe Bayazıt Subaşı, Cecilia Askham, Marco Cinelli; methodology: Ayşe Bayazıt Subaşı, Cecilia Askham, Marco Cinelli; formal analysis and investigation: Erlend Dancke Sandorf, Luis Cândido Dias, Danny Campbell; writing—original draft preparation: Ayşe Bayazıt Subaşı, Marco Cinelli, Cecilia Askham; writing—review and editing: Cecilia Askham, Erlend Dancke Sandorf, Luis Cândido Dias, Danny Campbell, Elçin Filiz Taş, Norihiro Itsubo, Christine Betty Nagawa, Christine Mugumya Kyarimpa, Mamadou Djerma, Bazoin Sylvain Raoul Bazie, Marco Cinelli; funding acquisition: Ayşe Bayazıt Subaşı, Cecilia Askham, Norihiro Itsubo; supervision: Marco Cinelli, Cecilia Askham, Elçin Filiz Taş.

Funding The work presented in this paper has been developed under the framework of the Global Guidance for Life Cycle Impact Assessment Indicators and Methods (GLAM) project of the Life Cycle Initiative, hosted by UNEP (www.lifecycleinitiative.org), and financially supported by the funding partners of the Life Cycle Initiative. Financial support for the survey company for Türkiye was provided by Ayşe Bayazıt Subaşı, PhD candidate at Istanbul Technical University. Financial support for the surveys in Burkina Faso and Uganda was received from NORSUS, Life Cycle Initiative hosted by UNEP (www.lifecycleinitiative.org), and EIFER. The surveys in Japan, India, and China were financially supported by JSPS KAKENHI Grant Numbers JP21H04944 (Grant-in-Aid for Scientific Research (A)). The authors would also like to acknowledge the advice from Associate Professor Kayo Murakami (Kwansei Gakuin University) on analytical and interpretive methods for the design of the survey.

Data availability Extensive data for this research is provided in the paper and in the Supplementary Information. Raw data from the responses to the survey are not publicly available to preserve individuals' privacy under the European General Data Protection Regulation.

Declarations

Conflict of interest The authors declare no competing interests.

References

Alriksson S, Öberg T (2008) Conjoint analysis for environmental evaluation: a review of methods and applications. Environ Sci Pollut R 15(3):244–257

Androulaki S, Psarras J (2016) Multicriteria decision support to evaluate potential long-term natural gas supply alternatives: the case of Greece. Eur J Oper Res 253:791–810

Bai S, Zhao X, Wang D, Zhang X, Ren N (2018) Engaging multiple weighting approaches and conjoint analysis to extend results acceptance of life cycle assessment in biological wastewater treatment technologies. Bioresource Technol 265:349–356

Bare JC, Hofstetter P, Pennington DW, de Haes HAU (2000) Midpoints versus endpoints: the sacrifices and benefits. Int J Life Cycle Ass 5:319–326

Ben-Akiva M, Steven RL (1985) Discrete choice analysis: theory and application to travel demand. MIT Press. https://0-discovery-ebsco-com.divit.library.itu.edu.tr/c/6k2lrh/viewer/pdf/7o57w7fet5

Bjørn A, Sim S, King H, Patouillard L, Margni M, Hauschild MZ, Ryberg M (2020) Life cycle assessment applying planetary and regional boundaries to the process level: a model case study. Int J Life Cycle Ass 25:2241–2254

Bos U, Horn R, Beck T, Lindner JP, Fischer M (2016) LANCA. Characterization factors for life cycle impact assessment, version 2.0. Fraunhofer Verlag, Stuttgart. https://doi.org/10.24406/publica-fhg-297633

BP (2022) Statistical review of world energy 2022. https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2022-full-report.pdf. Accesseed 13 Mar 2024

Castellan V, Benin L, Sala S, Pant R (2016) A distance-to-target weighting method for Europe 2020. Int J Life Cycle Ass 21(8):1159–1169

Cinelli M, Coles SR, Kirwan K (2014) Analysis of the potentials of multi criteria decision analysis methods to conduct sustainability assessment. Ecol Indic 46:138–148

Cinelli M, Spada M, Kadziński M, Miebs G, Burgherr P (2019) Advancing hazard assessment of energy accidents in the natural gas sector with rough set theory and decision rules. Energies 12:4178

Cinelli M, Kadziński M, Gonzalez M, Słowińsk R (2020) How to support the application of multiple criteria decision analysis? Let us start with a comprehensive taxonomy. Omega-Int J Manage S

Cinelli M, Kadziński M, Miebs G, Gonzalez M, Słowiński R (2022a) Recommending multiple criteria decision analysis methods with a new taxonomy-based decision support system. Eur J Oper Res 302:633–651

Cinelli M, Koffler K, Askham C, Amadei A, Arendt R, Bachmann TM, Barros B, Bjørn A, Dias LC, Laurent A, Motoshita M, SRupcic L, Sala S, Santos J, Scherer L, Steen B, Yokoi R (2022b) Criteria used to review weighting methods as part of the UN environment life cycle initiative's global guidance on environmental life cycle impact assessment indicators (GLAM) project. Paper presented at: SETAC Europe 32nd Annual Meeting; Copenhagen, Denmark

Cinelli M, Miebs G, Askham C, Amadei A, Arendt R, Bachmann TM, Bayazit Subasi A, Dias LC, Jolliet O, Koffler C, Laurent A, Motoshita M, Qian H, Rupic L, Santos J, Scherer L, Steen B (2023) A software for recommending weighting method(s) tailored to LCA studies. 11th International Conference on Industrial Ecology (ISIE2023)

De Laurentiis V, Secchi M, Bos U, Horn R, Laurent A, Sala S (2019) Soil quality index: exploring options for a comprehensive assessment of land use impacts in LCA. J Clean Prod 215:63–74

De Vlieghere A, van Veghel AS, Geeraerd A (2023) Life cycle assessment of importing canned tuna into Aruba through different supply chains, in varying can sizes and in oils, brine or tomato sauce. Int J Life Cycle Ass 28:1577–1589

Dias LC, Freire F, Geldermann J (2019) Perspectives on multi-criteria decision analysis and life-cycle assessment. In: Doumpos M, Figueira JR, Greco S, Zopounidis C (eds) New perspectives in multiple criteria decision making: innovative applications and case studies. Springer International Publishing, Cham, pp 315–329

Dias LC, Oliveira GD, Sarabando P (2021) Choice-based preference disaggregation concerning vehicle technologies. Cent Eur J Oper Res 29(1):177–200

Dias LC, Dias J, Ventura T, Rocha H, Ferreira B, Khouri L, MdoC L (2022) Learning target-based preferences through additive models: an application in radiotherapy treatment planning. Eur J Oper Res 302(1):270–279

Dias L, Morton A, Quigley J (2018) (Editors) Elicitation. The science and art of structuring judgement. Springer, Cham, 2018

Dias LC (2021) Sustainability assessment using the ELECTRE TRI multicriteria sorting method. Methods in sustainability science: assessment, prioritization, improvement, design and optimization, pp 197–214. https://doi.org/10.1016/B978-0-12-823987-2.00018-0

Doumpos M, Zopounidis C (2011) Preference disaggregation and statistical learning for multicriteria decision support: a review. Eur J Oper Res 209:203–214

Dyer JS, Sarin RK (1979) Measurable Multiattribute Value Functions. Oper Res 27(4):810–822

Eurostat (2023) Statistical Regions Level 2 in Turkey as of 18th July 2016 Level 1 Boundary Level 2 Boundary. https://ec.europa.eu/eurostat/documents/345175/7773495/TR.pdf. Accessed 26 Jul 2023

Goedkoop M, Spriensma R (2001) The Eco-indicator 99 a damage-oriented method for life cycle impact assessment methodology annex. www.pre.nl

Hélias A, Stanford-Clark C, Bach V (2023) A new impact pathway towards ecosystem quality in life cycle assessment: characterisation factors for fisheries. Int J Life Cycle Ass 28:367–379

Henriksen T, Astrup TF, Damgaard A (2020) Data representativeness in LCA: a framework for the systematic assessment of data quality relative to technology characteristics. J Ind Ecol 25(1):51–66

Hofstetter P (2000) Perspective in life cycle impact assessment: a structured approach to combine of the technosphere, ecosphere and valuesphere. Int J Life Cycle Ass 5:58–58

Horbach J, Rammer C, Rennings K (2012) Determinants of eco-innovations by type of environmental impact — the role of regulatory push/pull, technology push and market pull. Ecol Econ 78:112–122

Hoyos D (2010) The state of the art of environmental valuation with discrete choice experiments. Ecol Econ 69:1595–1603

Huijbregts MAJ, Steinmann ZJN, Elshout PMF, Stam G, Verones F, Vieira M, Zijp M, Hollander A, van Zelm R (2017) ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level. Int J Life Cycle Ass 22(2):138–147

Hüllermeier E, Słowiński R (2024) Preference learning and multiple criteria decision aiding: differences, commonalities, and synergies–part I. 4OR-Q J Oper Res. https://doi.org/10.1007/s10288-023-00560-6

Huppertz T, Weidema BP, Standaert S, De Caevel B, van Overbeke E (2019) The social cost of sub-soil resource use. Resources 8(1):19

Huppes G, Davidson MD, Kuyper J, van Oers L, Udo De Haes HA, Warringa G (2006) Eco-efficient environmental policy in oil and gas production in The Netherlands. Ecol Econ 61(1):43–51

IHME (2019) Global burden of disease (GBD). https://www.healthdata.org/research-analysis/gbd. Accessed 30 Mar 2023

ISO 14044 (2006) Environmental management-life cycle assessment-requirements and guidelines

Itsubo N (2015) Weighting. In: Hauschild MZ, Huijbregts MAJ (eds) Life cycle impact assessment. Springer, Netherlands, Dordrecht, pp 301–330

Itsubo N, Sakagami M, Washida T, Kokubu K, Inaba A (2004) Weighting across safeguard subjects for LCIA through the application of conjoint analysis. Int J Life Cycle Ass 9(3):196–205

Itsubo N, Sakagami M, Kuriyama K, Inaba A (2012) Statistical analysis for the development of national average weighting factors—visualization of the variability between each individual's environmental thoughts. Int J Life Cycle Ass 17:488–498

Itsubo N, Murakami K, Kuriyama K, Yoshida K, Tokimatsu K, Inaba A (2018) Development of weighting factors for G20 countries—explore the difference in environmental awareness between developed and emerging countries. Int J Life Cycle Ass 23(12):2311–2326

IUCN (2023) IUCN red list of threatened species. https://www.iucnredlist.org/. Accessed 27 June 2023

Jacquet-Lagrèze E, Siskos Y (2001) Preference disaggregation: 20 years of MCDA experience. Eur J Oper Res 130(2):233–245

Kadziński M, Martyn K, Cinelli M, Słowiński R, Corrente S, Greco S (2020) Preference disaggregation for multiple criteria sorting with partial monotonicity constraints: application to exposure management of nanomaterials. Int J Approx Reason 117:60–80

Keeney RL (1992) Value-focused thinking: a path to creative decision-making. Harvard University Press

Keeney RL, Raiffa H (1979) Decisions with multiple objectives: preferences and value trade-offs. IEEE Trans Syst Man Cybern B Cybern 9(7):403–403

Krantz D, Luce D, Suppes P, Tversky A (1971) Foundations of measurement, Vol. I: Additive and Polynomial Representations. https://philpapers.org/rec/KRAFOM

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

LCIn (2023) Global guidance for life cycle impact assessment indicators and methods (GLAM). https://www.lifecycleinitiative.org/activities/life-cycle-assessment-data-and-methods/global-guidance-for-life-cycle-impact-assessment-indicators-and-methods-glam/. Accessed 30 Mar 2023

Lippiatt BC (2007) Building for environmental and economic sustainability technical manual and user guide. https://www.nist.gov/publications/bees-40-building-environmental-and-economic-sustainability-technical-manual-and-user
Lombardi GV, Berni R, Rocchi B (2017) Environmental friendly food. Choice experiment to assess consumer's attitude toward "climate neutral" milk: the role of communication. J Clean Prod 142:257–262
Louviere JJ, Flynn TN, Carson RT (2010) Discrete choice experiments are not conjoint analysis. J Choice Model 3:57–72
Lueddeckens S, Saling P, Guenther E (2020) Temporal issues in life cycle assessment—a systematic review. Int J Life Cycle Ass 25:1385–1401
Mamine F, Fares Mh, Minviel JJ (2020) Contract design for adoption of agrienvironmental practices: a meta-analysis of discrete choice experiments. Ecol Econ
Manski CF (1977) The structure of random utility models. Theor Decis 8(3):229–254
Marttunen M, Belton V, Lienert J (2018) Are objectives hierarchy related biases observed in practice? A meta-analysis of environmental and energy applications of multi-criteria decision analysis. Eur J Oper Res 265(1):178–194
Matsatsinis NF, Grigoroudis E, Siskos E (2018) Disaggregation approach to value elicitation. Internat Ser Oper Res Management Sci 261:313–348
McFadden D (1974) Conditional logit analysis of qualitative choice behaviour. In: Frontiers in Econometrics, Academic Press, pp 105–152
Munda G (2016) Multiple criteria decision analysis and sustainable development. Multiple criteria decision analysis. Springer, New York LLC, pp 1235–1267
Murakami K, Itsubo N, Kuriyama K, Yoshida K, Tokimatsu K (2018) Development of weighting factors for G20 countries. Part 2: estimation of willingness to pay and annual global damage cost. Int J Life Cycle Ass 23(12):2349–2364
Pierrat E, Barbarossa V, Núñez M, Scherer L, Link A, Damiani M, Verones F, Dorber M (2023) Global water consumption impacts on riverine fish species richness in life cycle assessment. Sci Total Environ 854:158702
Qualtrics XM (2023) Qualtrics XM: The leading experience management software. https://www.qualtrics.com/uk/. Accessed 27 Jun 2023
Rakotonarivo OS, Schaafsma M, Hockley N (2016) A systematic review of the reliability and validity of discrete choice experiments in valuing non-market environmental goods. J Environ Manage 183:98–109
Richardson K, Steffen W, Lucht W, Bendtsen J, Cornell SE, Donges JF, Drüke M, Fetzer I, Bala G, von Bloh W, Feulner G, Fiedler S, Gerten D, Gleeson T, Hofmann M, Huiskamp W, Kummu M, Mohan C, Nogués-Bravo D, Petri S, Porkka M, Rahmstorf S, Schaphoff S, Thonicke K, Tobian A, Virkki V, Wang-Erlandsson L, Weber L, Rockström J (2023) Earth beyond six of nine planetary boundaries. Sci Adv 9:eadh2458
Rockström J, Steffen W, Noone K, Persson Å, Chapin FS, Lambin EF, Lenton TM, Scheffer M, Folke C, Schellnhuber HJ, Nykvist B, De Wit CA, Hughes T, Van Der Leeuw S, Rodhe H, Sörlin S, Snyder PK, Costanza R, Svedin U, Foley JA (2009) A safe operating space for humanity. Nature 461(7263):472–475
Rose JM, Collins AT, Bliemer M, Hensher DA (2018) NGENE (version 1.2.1). Choice Metrics
Sala S, Crenna E, Secchi M, Sanyé-Mengual E (2020) Environmental sustainability of European production and consumption assessed against planetary boundaries. J Environ Manage
Sala S, Cerutti AK (2018) Development of a weighting approach for the environmental footprint. Publications Office of the European Union, EUR 28562. https://doi.org/10.2760/945290. https://doi.org/10.2760/446145
Scarpa R, Rose JM (2008) Design efficiency for non-market valuation with choice modelling: how to measure it, what to report and why. Aust J Agr Res 52(3):253–282
Scherer L, Rosa F, Sun Z, Michelsen O, De Laurentiis V, Marques A, Pfister S, Verones F (2023) Kuipers KJJ (2023) Biodiversity impact assessment considering land use intensities and fragmentation. Environ Sci Technol 57:19612–19623
Siskos Y, Grigoroudis E, Matsatsinis NF (2005) UTA methods. International Series. Oper Res Manage Sci 78:297–343
Statista (2023) Education Worldwide – Statistics and Facts | Statista. https://www.statista.com/topics/7785/education-worldwide/#topicOverview. Accessed 17 Jul 2023
Steffen W, Richardson K, Rockström J, Cornell SE, Fetzer I, Bennett EM, Biggs R, Carpenter SR, de Vries W, de Wit CA, Folke C, Gerten D, Heinke J, Mace GM, Persson LM, Ramanathan V, Reyers B, Sörlin S (2015) Planetary boundaries: guiding human development on a changing planet. Science 347:1259855
Swait J, Louviere J (1993) The role of the scale parameter in the estimation and comparison of multinomial logit models. J Market Res 30(3):305–314
Tervonen T, Sepehr A, Kadziński M (2015) A multi-criteria inference approach for anti-desertification management. J Environ Manage 162:9–19
UNEP (2023) The Life Cycle Initiative | UNEP - UN Environment Programme. https://www.unep.org/explore-topics/resource-efficiency/what-we-do/life-cycle-initiative. Accessed 11 Apr 2023
UNEP-GLAM (2021) Defining the plan for a global life cycle impact assessment method (GLAM) - life cycle initiative. https://www.lifecycleinitiative.org/defining-the-plan-for-a-global-life-cycle-impact-assessment-method/
UNEP LCI (2021) Global LCIA guidance (GLAM) phase 3 scoping document global LCIA guidance phase 3 'creation of a global life cycle impact assessment method'. https://www.lifecycleinitiative.org/wp-content/uploads/2021/02/GLAM3-Scoping-document.pdf
U.S. Geological Survey (2022) Mineral commodity summaries 2022: U.S. Geological Survey, 202. https://doi.org/10.3133/mcs2022
Verones F, Bare J, Bulle C, Frischknecht R, Hauschild M, Hellweg S, Henderson A, Jolliet O, Laurent A, Liao X, Lindner JP, Maia de Souza D, Michelsen O, Patouillard L, Pfister S, Posthuma L, Prado V, Ridoutt B, Rosenbaum RK, Sala S, Ugaya C, Vieira M, Fantke P (2017) LCIA framework and cross-cutting issues guidance within the UNEP-SETAC life cycle initiative. J Clean Prod 161:957–967
Victor P, Hanna S, Kubursi A (1998) How strong is weak sustainability? In: S. Faucheux, M. O'Connor and J. Straaten, eds. Sustainable development: concepts, rationalities and strategies: Springer Netherlands; 13:195–210
World Bank (2023a) GDP (Current US$) | Data. https://data.worldbank.org/indicator/NY.GDP.MKTP.CD. Accessed 13 Apr 2023
World Bank (2023b) GDP per Capita (Current US$) | Data. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD. Accessed 27 Jun 2023
World Bank (2023c) World Bank Country and Lending Groups – World Bank Data Help Desk. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups. Accessed 30 Aug 2023
World Bank (2023d) Individuals using the Internet (% of Population) | Data. https://data.worldbank.org/indicator/IT.NET.USER.ZS. Accessed 30 Aug 2023
World Bank (2023e) Population Estimates and Projections | Data Catalogue. https://datacatalog.worldbank.org/search/dataset/0037655/Population-Estimates-and-Projections. Accesseed 17 Jul 2023
Zerva A, Grigoroudis E, Karasmanaki E, Tsantopoulos G (2021) Multiple criteria analysis of citizens' information and trust in climate change actions. Environ Dev Sustain 23:7706–7727

Zheng J, Lienert J (2018) Stakeholder interviews with two MAVT preference elicitation philosophies in a Swiss water infrastructure decision: aggregation using SWING-weighting and disaggregation using UTAGMS. Eur J Oper Res 267:273–287

Authors and Affiliations

Ayşe Bayazıt Subaşı[1] · Cecilia Askham[2] · Erlend Dancke Sandorf[3] · Luis Cândido Dias[4] · Danny Campbell[5] · Elçin Filiz Taş[1] · Norihiro Itsubo[6] · Christine Betty Nagawa[7] · Christine Mugumya Kyarimpa[8] · Mamadou Djerma[9] · Bazoin Sylvain Raoul Bazie[9,10] · Marco Cinelli[11,12]

✉ Marco Cinelli
m.cinelli@luc.leidenuniv.nl

1 Department of Architecture, İstanbul Technical University, Taşkışla Cad. No:2 Harbiye, Şişli, Istanbul, Turkey

2 NORSUS, Norwegian Institute for Sustainability Research, Stadion 4, 1671 Kråkerøy, Norway

3 School of Economics and Business, Norwegian University of Life Sciences, Christian Magnus Falsens Vei 30, 1433 Ås, Norway

4 CeBER, Faculty of Economics, University of Coimbra, Av. Dias da Silva 165, 3004-512 Coimbra, Portugal

5 Economic Division of Management School, University of Stirling, Stirling FK9 4LLA, Scotland, UK

6 School of Creative Science and Engineering, Waseda University, 1-104 Totsukamachi, Shinjuku-Ku, Tokyo 169-8050, Japan

7 Department of Forestry, Biodiversity and Tourism, College of Agricultural and Environmental Sciences, Makerere University, P.O. Box 7062, Kampala, Uganda

8 Department of Chemistry, Faculty of Science, Kyambogo University, P.O. Box 1 Kyambogo, Kampala, Uganda

9 Sustainable Water and Environment Expertise and Technologies (SWEET), Avenue de la Dignité, rue 17.206 Ouagadougou, Burkina Faso

10 Centre Universitaire Polytechnique de Manga/Université Norbert Zongo, BP 376 Koudougou, Burkina Faso

11 Decision Engineering for Sustainability and Resilience Laboratory, Leiden University College, Faculty Governance and Global Affairs, Leiden University, Anna van Buerenplein 301, 2595 DG, The Hague, The Netherlands

12 Institute of Environmental Sciences (CML), Faculty of Science, Leiden University, P.O. Box 9518, 2300 RA, Leiden, The Netherlands