Agricultural Flood Damage Assessment: A Systematic Review of Methods, Coverage, and Integration Gaps

report's referee 2

2026-11-16

Table of contents

1 Summary and recommendation **1**

2 Major comments **2**

2.1 Major articles published in NHESS were not identified and they contradict the gaps claimed in the manuscript . . . 2

2.2 The article selection does not adequately cover the literature, and the inclusion-exclusion is not clear . . . 3

2.3 The coding behind the counts is not provided, the counts are not verifiable . . 4

2.4 Damage categories and assessment methods are confused, and the typology is not positioned relative to the reference frameworks . . . 4

2.5 The "post-2013 frontiers" are not new and are weakly supported in the review . 5

2.6 The use of ex-ante and ex-post evaluation methods is conflated . . . 5

3 Recommendations and requested actions **5**

1 Summary and recommendation

The manuscript reviews 65 journal articles (1984–2025) on the economic assessment of flood damage to agriculture. It proposes a Methods-by-Damage-Category Matrix (MDCM) that maps assessment methods onto damage categories, and presents the work as an update of the review by Brémond et al. (2013). It adopts the damage categories of that review, adds three further categories and concludes that current estimates largely capture only direct damage.

The topic is suitable for NHESS and an update of Brémond et al. (2013) is welcome. The effort to assemble and classify such a corpus is appreciated.

In its current form, however, the main findings are not treated rigorously enough for reasons outlined below:

1. major articles on the same topic, four of them published in NHESS, were not identified, and they address several gaps that the manuscript presents as unresolved (Brémond et al., 2022; Molinari et al., 2019; Nortes Martínez et al., 2021; Bill-Weilandt et al., 2026; and Scorzini et al., 2021 in the International Journal of Disaster Risk Reduction);
2. the processus of selection of articles needs to be checked and explained to understand these omissions;
3. the coding at the article level that underlies the percentages is not provided and the counts cannot be verified;
4. The manuscript does not sufficiently distinguish damage categories from assessment methods and does not rely enough on the reference frameworks of Merz et al. (2010), Meyer et al. (2013) and Brémond et al. (2013), which leads it to present already-established damage categories as new; the "post-2013" categories presented as new are already present in Merz et al. (2010) and Brémond et al. (2013);
5. ex-ante and ex-post uses of damage assessment are not sufficiently distinguished. For these reasons, my recommendation is a major revision.

2 Major comments

2.1 Major articles published in NHESS were not identified and they contradict the gaps claimed in the manuscript

Five papers that address exactly this topic—four of them published in this journal—are neither cited nor included in the 65-study corpus. Three are methodological contributions (Brémond et al., 2022; Molinari et al., 2019; Nortes Martínez et al., 2021); another is an application of one of those methods (Scorzini et al., 2021, International Journal of Disaster Risk Reduction); the fifth, published in 2026, provides a review of rice-damage models. These works must be taken into account when revising the "gaps" statements identified by the manuscript.

These articles are relevant to at least six statements made in the manuscript:

1. **"No established protocol exists for converting physical damage signals into economic loss"** (lines 73-74; conclusion line 639). Process-based models such as **floodam.agri** (Bremond et al, 2022) and **AGRIDE-c** (Molinari et al, 2019), as well as several operational methods cited in Brémond et al. (2013), already provide such physical-to-economic conversion chains.

2. **"No systematic protocol for adjusting damage functions across agronomic contexts"** (lines 215-217), **"no coherent framework for matching method choice to data context"** (lines 577-579), and recommendation 5.3(c) concerning the reporting of data requirements and transfer conditions. **Brémond et al. (2022)** operationalises precisely these aspects, and **Molinari et al. (2019)** documents a transfer of a damage function. **Scorzini et al. (2021)** and **Bill-Weilandt et al. (2026)** also treat this question.

3. **Perennial crops "require a fundamentally different framework"** (lines 379-383). Both **Brémond et al. (2022)** and **Scorzini (2021)** explicitly model the multi-year income stream associated with perennial systems. **Nortes Martínez et al. (2021)**, who model flood damage on wine cooperative systems also take into account the multi-year loss associated with perennial systems.

4. **"First" systematic characterisation and "no prior review"** (lines 75-76, 231-236). **Molinari et al. (2019)**, **Brémond et al. (2022)** and **Bill-Weilandt et al. (2026)** constitute partial syntheses on the same dimensions (hazard parameters, monetary indicator, validation, transferability).

5. **Method-induced uncertainty "rarely reported"** (lines 224-228). **Molinari et al. (2019)** and **Scorzini et al. (2021)** address this.

6. **Supply-chain and farm-scale propagation has "no established assessment method"** (lines 175-177): **Nortes Martínez et al. (2021)** address this question.

2.2 The article selection does not adequately cover the literature, and the inclusion-exclusion is not clear

The omission of the five papers mentioned above—four of which were published in NHESS and all of which clearly fall within the declared scope—raises questions about how the articles were selected, beyond the simple issue of the search string itself. The manuscript states that Rayyan's AI-assisted feature was used as a prioritisation tool, that all inclusion and exclusion decisions were made by human reviewers, and that no record was excluded on the basis of the AI ranking alone (lines 265-271). However, the flow diagram shows that the reduction from 3 881 articles to 84 full-text articles was achieved by applying two keyword filters ("flood" AND "agriculture", then "economic") to titles and abstracts, which removed 3 627 and 170 records respectively—that is 3 797 out of 3 881 (98 %). The manuscript describes human reading only from the 84 full-texts onward; for the 98 % of records eliminated by the filters, no human screening of titles and abstracts is mentioned.

In a systematic review conducted with Rayyan, human screening of titles and abstracts—ideally by two reviewers with conflicts resolved—is the step that guards against omissions of this kind; here it appears to have been replaced by a keyword rule. I do not suspect bad faith, but it seems that the corpus cannot be regarded as the product of a truly "human-screened"

systematic selection, and that the statements regarding literature coverage, the persistence of gaps over four decades, and the number of studies per category all rest on a selection that has not been validated.

Therefore I would ask the authors to:

- Specify how many reviewers screened titles and abstracts, at which stages, and how many records each reviewer actually read;
- Deposit the Rayyan export containing the decision, reviewer, and conflict status for every record;
- Compare the results of the pipeline against the known-item set (the articles compiled in Brémond et al. 2013 and the five papers listed above) and indicate at which stage each of those items was lost;
- Have the titles and abstracts of the records eliminated by the keyword filters screened by humans, or alternatively restrict the claims to what the current selection can actually support.

2.3 The coding behind the counts is not provided, the counts are not verifiable

The manuscript (lines 285 and 660) states that Supplement A contains the damage-category classifications. In fact, it is only a bibliography. None of the percentages by damage category reported in the paper can be verified.

The authors should provide a supplementary table that, for each study, lists: the country, crop, flood type, the method(s) used, the damage categories covered, the data employed, the monetisation indicator, the validation procedure, and the uncertainty treatment, together with the coding rules. All percentages should then be recomputed from this table.

2.4 Damage categories and assessment methods are confused, and the typology is not positioned relative to the reference frameworks

The risks is to add confusion to a literature that has worked to harmonise its terminology, rather than clarifying it.

- The categories presented in the article are not new and are misreferenced. The manuscript states that it starts from a three-axis taxonomy "introduced by Brémond et al. (2013)." The "direct/indirect" and "tangible/intangible" typologies already existed in the risk-economics literature (Merz et al., 2010; Meyer et al., 2013). Brémond et al. (2013) added the immediate/induced distinction. Merz et al. (2010) is cited only for the definition of risk even though they had already proposed a typology of flood-related damages, and Meyer et al. (2013) is not cited at all.

- Reference frameworks distinguish damage categories from evaluation methods. Meyer et al. (2013) define five cost categories, including business interruption, because each requires different methods. Brémond et al. (2013) also first define damage categories and then review the methods for assessing them.

2.5 The "post-2013 frontiers" are not new and are weakly supported in the review

- Brémond et al. (2013) already addressed issues of induced damage to soil, multi-year losses, farmers' access to external resources—including cash —as well as the "increased risk of bankruptcy" that requires a farm-scale approach. Salinity and contamination are also mentioned as unmodeled processes in Brémond et al. (2022), and AGRIDE-c cites long-term soil productivity as a limitation. The emphasis on a "post-2013" framing is therefore exaggerated.
- **The categories presented as new were already included in Merz et al. (2010).** That review lists psychological impacts and negative effects on ecosystems as direct intangible damage, and trauma and loss of trust as indirect intangible damage; it classifies business interruption within the flooded area, agricultural soil erosion, crop destruction, and livestock damage as direct tangible damages; it describes soil damage in agriculture as pollutant deposition and loss of soil structure due to compaction or erosion (Section 3.5.4); it notes that indirect effects can persist for months or even years (Section 2.1); and, citing Pfurtscheller and Schwarze (2010), it includes the availability—or lack—of financing and reconstruction aid among the factors governing indirect damage (Section 4). Thus, psychological and ecosystem damages, soil compaction, and the role of financing are not categories identified for the first time by the present review.

2.6 The use of ex-ante and ex-post evaluation methods is conflated

The introduction frames the issues through benefit-cost analysis, insurance, and compensation. However, these three approaches require different inputs. Benefit-cost analysis needs an ex-ante estimate of damages as a function of hazard parameters in order to compute expected annual damages for various scenarios, whereas compensation and some insurance schemes can rely on ex-post assessments. Distinguishing between these two approaches would be useful.

3 Recommendations and requested actions

The necessary revisions involve revisiting the methodology, the corpus selection, and the coding, then reformulating the results and arguments accordingly. I would be happy to review a revised version.

Revisions to be made

1. Validate the selection procedures using a known-item recall test (Brémond et al. 2013, Molinari et al. 2019, Brémond et al. 2022, Bill-Weilandt et al. 2026): use a wider keyword filter, reinstate human screening of titles and abstracts, deposit the per-record decisions in Rayyan, justify language and gray-literature limits, and update the search through 2026.
2. Provide the extraction table at the article level: present the coding rules and recompute every count and percentage in the abstract, main text, tables, and conclusion.
3. Clarify the role of AI tools. The manuscript mentions AI-assisted prioritisation in Rayyan. Given the questions raised above about the selection of articles and the coding, I would invite the authors to state clearly whether and how AI tools were used at each stage (screening, data extraction, classification and drafting), and to describe how the results were checked by the authors.
4. Use and cite the typology on the frameworks of Merz et al. (2010), Meyer et al. (2013) and Brémond et al. (2013): keep damage categories and assessment methods strictly separate, distinguish modeling strategy, data source, scale, ex-ante/ex-post use and validation status; include process-based and expert-knowledge models.
5. Incorporate the works of Brémond et al. (2022), Molinari et al. (2019), Scorzini et al. (2021), Nortes Martínez et al. (2021) and Bill-Weilandt et al. (2026): revisit the statements regarding identified gaps, transferability frameworks, perennial crops, uncertainty and novelty.
6. Correct the misattributed and unsupported statements listed in the manuscript.