Response to Reviewer 2

How to cite. Ubeda-Trujillo, I., Werner, M., Bertini, C., Coenders-Gerrits, M., and Jewitt, G.: Detection of agricultural flash droughts using impact-informed thresholds, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2026-1010, 2026.

We thank the Reviewer for the detailed questions, insightful comments, and time invested in reviewing the manuscript. These suggestions have helped us improve the clarity of the paper. The reviewer's comments and the authors' responses are provided below, with reviewer comments shown in green and responses shown in blue.

General comments

The manuscript presents a modified approach for detecting Agricultural Flash Droughts by calibrating the Standardized Evaporative Stress Ratio (SESR) with crop management periods in various soil textures. This research significantly contributes to the advancement of monitoring droughts for improved decision-making in agricultural contexts. However, the document requires major revision and language editing before it can be accepted for publication.

We thank the reviewer for the comment. The language will be revised throughout the manuscript.

Specific comments

1. The abstract is complete but consider starting the main results section with your approach instead of beginning with the literature-based approach.

We thank the reviewer for the comment. We agree with the reviewer that the abstract placed greater emphasis on the limitations of literature-based thresholds rather than on the performance of the proposed approach. To address this, we have revised how the results are presented in the abstract to more explicitly highlight the performance of the calibrated impact-informed thresholds relative to the literature-based thresholds.

2. It is unclear how to assess the potential agricultural relevance of the SESR for detecting anomalous evaporative stress during years with drought -induced crop losses. What method or technique should be used for this assessment?

We thank the reviewer for this observation. The assessment was conducted using Pearson correlation and linear regression analyses (statistical techniques), comparative distribution analysis (overarching method), and non-parametric statistical tests (Mann–Whitney U and Kolmogorov–Smirnov tests; statistical techniques).

However, we would first like to clarify that the purpose of this section was not to assess the agricultural relevance of SESR as an indicator. Rather, the objective was to assess whether evaporative stress, represented by SESR, responded differently between years with and without drought-induced crop losses under the dryland conditions of the study area. This assessment was conducted using the methods and techniques already presented throughout this section and mentioned above. Therefore, to avoid this misunderstanding, we will revise the objective and rename the section accordingly, while retaining the same methods and techniques mentioned.

To better reflect in the result section this revised objective, we propose retaining Figure 5 as the primary figure, as it more directly supports the objective and Figures 4 and S2 will be moved to the Supplementary Material.

As explained in our response to Reviewer #1 (Comment #9). This objective was particularly important because, in dryland environments, high evaporative demand and chronically low soil moisture make evaporative stress part of the normal background conditions. Under these conditions, it cannot be assumed that literature-based SESR thresholds consistently distinguish years with and without drought-induced crop losses across crop management periods and soil textures. Therefore, before calibrating impact-informed SESR thresholds, it was necessary to assess whether SESR differentiated years with and without drought-induced crop losses impact. If no such relationship existed, the subsequent calibration of SESR thresholds against agricultural impact benchmarks would have lacked a sufficient basis.

To conduct the assessment, reference years with and without drought-induced crop losses were first identified using the VHI-based benchmark described in Sect. 2.3.1. Using these reference years, SESR time series were extracted from February to July at 8-day resolution for each sampled location and grouped by soil texture class. Then, Pearson correlation and linear regression were used as the initial statistical techniques to analyse the relationship between evaporative stress signals represented by SESR and vegetation conditions represented by VHI. Additionally, boxplots were used visual technique to contrast the behaviour of SESR between years with and without crop losses. This allowed for the simultaneous evaluation of shifts in central tendency (the median), dispersion (the interquartile range), and overall data symmetry, providing a clear visual assessment of how evaporative stress distributions responded to drought conditions before conducting the formal significance tests. Finally, Mann–Whitney U and Kolmogorov–Smirnov tests were the specific non-parametric techniques applied to evaluate the statistical significance of the differences between those distributions.

3. The methods section is somewhat challenging to follow. It should be organized according to the study´s objectives.

We thank the reviewer for this comment. This comment was also raised by Reviewer #1. As explained in our response to Reviewer #1 (Comment #3), the revised manuscript will rename and reorganise the method sections to align directly with the study sub-objectives, thereby improving the overall flow and consistency between the objectives, methods, and results.

4. Results should be presented in accordance with the study´s objectives. Therefore, section 3.3 is a part of objective 2 (section 3.2, uncertainty of SESR threshold).

We thank the reviewer for this comment. This comment was also raised by Reviewer #1. As explained in our response to Reviewer #1 (Comment #3), the revised manuscript will rename and reorganise the results sections to align directly with the study objectives and methods.

5. The label on Figure 8 mentions black cells, but none of the cells are actually black. What is black is the outline of some cells.

We thank the reviewer for this observation. Indeed, the 'black cell' mentioned in lines 497–498 refers to the cells distinguished by their black outline/border. In the revised version of the manuscript, we will change 'black cell' to 'black-bordered cell' (or 'cell with a black outline') to accurately describe the visual representation in Figure 8.

6. In section 2.3.6 what are you referring to with validation metrics? how does the procedure work by using these two metrics?

We thank the reviewer for this comment. This comment was also raised by Reviewer #1. As explained in our response to Reviewer #1 (Comments #7 and #8). The literature-based approach refers to the original SESR flash drought detection method of Christian et al., 2019, in which the drought onset threshold is defined using a fixed SESR percentile and is not calibrated against local agricultural drought impacts. This study does not introduce a new flash drought detection. Instead, it retains the SESR-based flash drought method of Christian et al., 2019, including the rapid intensification criteria, and modifies only the drought onset threshold criterion. Therefore, V1 and V2 are binary validation metrics used to assess whether events detected using the fixed or calibrated SESR drought onset thresholds together with the same rapid intensification criteria coincided with the two agricultural impact benchmarks. A value of 1 indicates coincidence between a detected event and the corresponding benchmark (hit), whereas a value of 0 indicates no coincidence (miss). Specifically, V1 = 1 indicates coincidence with the VHI-derived agricultural impact benchmark, whereas V2 = 1 indicates coincidence with documented drought impacts reported in the benchmark literature.

Because the study compares the same detection method under two different drought onset threshold parameterisations, improvement was assessed by comparing the agreement of detected events with the two agricultural impact benchmarks. The calibrated SESR threshold was considered to improve performance when it showed greater agreement with the agricultural impact benchmarks than the original literature-based threshold.

7. There is no mention of the validation of the literature-based thresholds in the discussion.

We thank the reviewer for the comment. We would like to clarify that the validation of literature-based thresholds is already addressed in the Discussion section (Pages 17-18, Lines 540-543, and 558). We discuss how literature-based thresholds showed limited skill compared to our calibrated ones. As specified also in our previous answer, the literature-based threshold is validated against agricultural impact benchmarks with the metrics V1 and V1. As well as the validation is presented in the results section 3.4.

Technical corrections

Line 91-97: The idea is too long; it should be split into two or three connected sentences.

The sentence will be split into two or three connected sentences.

Line 125: remove capital letter in “coverS”

This capital letter will be changed.

Line 252 insert the letter “d” in “characterized”

The letter “d” will be added.

Line 265 use modifies instead of refines

The word “modifies” will be used instead of “refines”.

Line 317 "…values were then created: one for years with drought-induced crop losses, and one for years without."

This sentence will be corrected.

Line 367: Why do you separate precision from the group of performance metrics?

We thank the reviewer for this question. Precision was mentioned separately in this sentence simply because it had already been introduced and addressed in detail at the beginning of this section. Therefore, the phrasing "together with the precision metric" was used as a transition to connect the previously discussed metric with the newly introduced performance and uncertainty metrics, rather than to exclude it from the overall group.

Line 531-537: The idea is too long; it should be split into two or three connected sentences.

This will be corrected.

Line 539-542: The idea is too long; it should be split into two connected sentences.

The idea will be split into two or three connected sentences.

Line 542-547: The idea is too long; it should be split into three connected sentences.

The idea will be split into two or three connected sentences.

Line 554-557: The idea is too long; it should be split into three connected sentences.

The idea will be split into two or three connected sentences.

Line 580-584: The idea is too long; it should be split into two connected sentences.

The idea will be split into two or three connected sentences.