Dear Joris,

Thank you for taking the time to review our manuscript and for the valuable comments and suggestions.

We will provide a point-by-point reply to all your comments in our final response letter, but we would like to take the opportunity address your main concerns while the interactive discussion is open.

Below your comments are displayed in orange, and the author response in blue:

Referee comment:

The objectives should be better defined. The first objective focusses on testing of the model's capabilities. This does not seem to be too ambitious. No matter what is the outcome, this objective will always be achieved. So please refine this objective to make it more ambitious. The second objective seems related to the GLUE approach, I suggest to explicitly include the GLUE approach in this objective. The last objective is similarly not too ambitious (either testing or analysing something will always be achieved).

Author's response:

This is a very good point – thanks for bringing it up.

We will refine the phrasing of our objectives to highlight the ambition and importance of our work. That is, our study employs a rejectionist limits-of-acceptability approach within the GLUE framework to test a widely used soil erosion and sediment delivery model in settings where it is routinely applied but not well evaluated against field data (particularly not high-quality, long-term monitoring data). In this context, "testing capability" is not a passive task where success is guaranteed, but a rigorous attempt at model falsification (Beven, 2006), which is critical to improve models and understanding, but rarely done in environmental modelling in general and erosion modelling in particular.

Referee comment:

The concept of aggregating the data using different spatiotemporal resolutions has not been mentioned in the Introduction. I was expecting that this would go in some direction of using different spatial and temporal resolutions (different cell sizes and time steps, for instance). However, this is totally not the case. The authors instead use the long-term median model outcome, instead of the annual outcomes (the way USLE-type of models should actually be used). And the spatial aggregation is related to the two different conservation types considered. I'm not sure if this requires to be included in an objective.

Author's response:

Apologies if this topic was not sufficiently developed in the introduction – we will make amends in the revised manuscript. There might be some misunderstandings we'd like to clarify:

(i) Our analysis does compare two different temporal resolutions (perhaps we need to check the phrasing to make this more precise). We evaluate the model on annual timesteps (from 1994 to 2001) and then on the aggregated eight-year median. However, it is not correct that USLE-type models were designed to operate with an annual resolution. The original USLE was developed to predict long-term average annual soil loss, not to produce annual model outputs (Wischmeier and Smith (1978) - Page 2: "The USLE is an erosion model designed to predict the longtime average soil

losses in runoff from specific field areas in specified cropping and management systems."; or page 40: "The USLE is designed to predict longtime-average soil losses for specified conditions."). Our finding that the model performs better when aggregated over the eight-year study period is therefore fully consistent with the original design and purpose of the USLE (or ABAG in our case).

(ii) We apologise if our spatial aggregation (single watersheds or combined as field- or structure-dominated) was not clear or gave you the impression we looked at different cell-sizes for the input data. The aggregation is not simply between two different conservation types. It is between: (i) Field-dominated watersheds (W01-W04), which are characterized by in-field soil conservation practices (e.g., no-till, crop rotation), retention ponds in W01 and W02 and small grass-strips. (ii) Structure-dominated watersheds (W05, W06), which include in-field soil conservation practices plus additional linear landscape structures (specifically, retention ponds and a big, grassed waterway) designed to trap sediment. This distinction is important for our two-phase GLUE approach, where we first condition in-field parameters on the field-dominated group (Phase 1) and then use those to condition the structural parameters on the structure-dominated group (Phase 2). This aggregation is also important to evaluate the transferability of model conditioning between watersheds. We will make these points clearer in the revised manuscript.

References

Beven, K.: A manifesto for the equifinality thesis, J HYDROL, 320, 18-36, 10.1016/j.jhydrol.2005.07.007, 2006.

Wischmeier, W. H. and Smith, D. D.: Predicting rainfall erosion losses: A guide to conservation planning, Agriculture Handbook, 537, Department of Agriculture, Science and Education Administration, United States, 65 pp.1978.