**Authors' response letter** (hess-2023-304)

The following are our responses to comments by the community

Dear Dr Oluwafemi Adeyeri,

We thank you very much for your constructive comments and suggestions. The comments are very useful to improve the quality of our paper. All concerns and suggestions raised are addressed accordingly. Please find below your comments and our responses. Note that the authors' responses are highlighted in red.

1. P3L25-33: These lines are unclear. It is unclear if the authors are evaluating what you already stated as a setback. Are you inheriting these setbacks into your focus?

Response #1: It is true that there are challenges with accurately representing soil physical properties in land surface models, our focus is to investigate how much this uncertainty would impact the simulated land state conditions and whether using a state-of-the-art soil database such as SOILGRID with PTF would add any value.

2. P5L1: Does this improvement assume no heat flux transfer between soil layers? How is this practicable?

Response #2: The use of a zero thermal boundary condition at the bottom deep soil is common in land surface models and is an option in the NOAH-MP model. It assumes that the thermal gradient is small and does not influence temperature much at the shallow soil depth. But we have used a different physics option provided in the model (see Table 3 of the paper) to consider the thermal regime at the subsurface, in our study.

3. P6L29: I am not sure landuse is static, as notable transformations could occur in 10 years.

Response #3: We agree with your point of view, the static here means it does not change over the course of a year. We will revise this to avoid confusion.

4. P5L3: Does the model account for scale-gap effects considering input data resolutions?

Response #4: Yes, the WRF preprocessing system (WPS) was used to harmonise and generate the geographic files and forcing files at 1 km grid space.

5. P6L10: What are the bases for the classification? What is the method used and what is the rationale behind it?

Response #5: Thanks. The goal was to use a high resolution (e.g. 100 m Corine) input land cover product with 44 classes that better represent Ireland's landscape as inputs. The NOAH-MP model provides a

lookup table of vegetation properties that are linked to only 21 land cover categories based on MODIS land cover, reclassifying Corine to match the MODIS was necessary to easily link the 21 vegetation properties. We have used a simple method that merges the common or almost similar land cover types in terms of attributes to reduce the number of categories.

6. P6L334: How is the stability of the model ascertained?

Response #6: Thanks. This was determined when the difference between present and previous spin-up soil moisture outputs is small or negligible ( ~ $10^{-3}$). By using the input climatology, this was quickly achieved after 3-5 spin up runs.

7. P7L4: Are there any uncertainties regarding the station data quality and how are thgey assessed?

Response #7: Thanks. Assessing uncertainties in the station data is beyond the scope of this work; however, we have ensured minimum measurement standards in our Terrain-AI stations, including calibration of TDR sensors.

We will revise the paper accordingly to properly acknowledge the potential uncertainties due to different limitations. These will include spatial scale mismatches between point-based observations and model 1 km grids; differences in soil depths between observations and model; uncertainties in soil moisture measurements/satellite products.

8. P7L14: What does time for the soil to settle mean?

Response #8: In the process of installing the TDR sensors, including digging and covering holes or trenches where the sensors are buried, the soil-to-sensor contact must be good, with no air gaps. Depending on the soil condition, time is often required after installation to achieve this or for the soil to properly settle around the sensor, ideally reaching saturation.

9. P7L34: Is this a standard product or the authors develop it? If it is a standard product, it needs referencing. If developed by authors, what method was used for this fusion, bearing in mind the transfer of error from fusion approaches?

Response #9: It's a global satellite product, not developed by the authors. Relevant references have been provided and the link to access the product is also provided in the 'code and data availability' section

10. P8L1: Is this different from previously explained?

Response #10: There are different soil moisture products from ASCAT, the statement here explicitly states that we have used the 1 km daily ASCAT SWI which covers the European region in our study.

11. P8L11: what does extracted at model resolution mean and how was this done?

Response #11: What we mean here is that we extracted modeled values for the stations' colocated 1 km by 1 km grid cells. This was done using the 'terra' R package, by converting the stations' geometry into spatial vector points, then using this to extract the corresponding model grid values.

12. P9L27: Seasonal variability can not be assessed simply from the time series presented. In fact, 4f and g are misleading as they did not show seasonal variabilities.

Response #12: Thanks. This will be revised accordingly.

13. P9L35: This assumption does not warrant generalization. The uncertainty associated with comparing the model areal grid to measurement points should be quantified.

Response #13: Quantifying uncertainty associated with scale mismatch between points and grid is beyond the scope of the current work. However, as the standard practice in model evaluation, we have acknowledged this limitation in P8L18-20.

We will further revise the paper accordingly to properly acknowledge the potential uncertainties due to different limitations. These will include spatial scale mismatches between point-based observations and model 1 km grids; differences in soil depths between observations and model; uncertainties in soil moisture measurements/satellite products.

14. P10L1: I suggest the authors combine error metrics instead of individual ones. Since error metrics are sensitive to differences in precision, it is vital to have a combined metric to account for the different statistical properties of an ideal model performance. Collectively assessing these statistical metrics provides a comprehensive understanding of the performance of each model. This approach provides valuable insights for the overall evaluation of the models. The ensemble method of statistical metrics will further reveal the overall efficacy and reliability of the models.

Response #14: Thanks for your suggestion. We are concerned about not introducing confusion, especially in relation to interpreting a weighted metric. Each metric tells something different about our results, understanding these differences is important for the overall interpretation of the results.

15. Figure 5: It is also important to understand this distribution's mean changes. Also, check if these changes are significant.

Response #15: Thanks. This point has been addressed already in the boxplots presented in Figure 5.

16. Figures 10 and 11: No section in the methodology explicitly addressed how these were calculated or generated.

Response #16: We do not agree with this point. P9L15-22 and Table 4 explicitly explain how the maps of drought categories were generated.

17. Address potential limitations of your study, uncertainties in the results, and sources of error. Acknowledge the challenges and potential biases that might affect the interpretation of the findings.

Response #17: Thanks for your suggestion. This concern will be addressed properly in the paper, as already stated in response #13 above.

18. The manuscript's introduction appears overly extensive, encompassing information tangential to the study objectives. A thorough revision to streamline the introduction and enhance focus is recommended.

Response #18: We do not agree with this point. We believe that the current introduction is concise and provides detailed background and existing knowledge (in a brief literature review) about the topic.