Comment on essd-2021-420
Anonymous Referee #1

Referee comment on "A Central Asia hydrologic monitoring dataset for food and water security applications in Afghanistan" by Amy McNally et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2021-420-RC1, 2022

Summary: the authors provide a data description, summary of applications, and application case study for the Famine Early Warning System Land Data Assimilation System (FLDAS) for Central Asia. The discussion paper succeeds as a data description. I found the writing to be clear, the explanation of methods reasonably complete, and the presentation of evaluation data to be adequate for purposes of data description. The description of applications is a particular strength of this paper, as it is evident that the FLDAS is actively used in this region, and it is interesting to read of the specific La Nina drought example in which FLDAS products were leveraged for food security analysis.

I believe that the discussion paper can be accepted as a final paper in Earth System Science Data with minimal modification. I do have the following questions for the authors to consider.

- It would be useful for the modeling community to understand a bit more about the data choices summarized in Table 1. Why were each of these parameter datasets selected? For example, the FAO soil texture dataset has been replaced by ISRIC in many applications, and the NCEP vegetation fraction dataset is a low resolution climatology. I don’t ask that the authors change these settings, but given the status of FLDAS as an operational LDAS system with global scope I expect that the authors are in good position to provide readers with some guidance regarding the choice of parameter sets. (Also: a minor note on Table 1—the row labels could be improved: why is the FAO soil data simply listed as “parameters” and the snow albedo simply listed as “albedo”? It looks like words were dropped.)
- Section 2.3: The authors explain how they used the different precipitation datasets, but I’m confused about the rationale. IMERG is not a particularly new product—it’s been available for several years. Why does the system still use GDAS precipitation and only include IMERG as a data comparison? The IMERG Late (or Early) runs would have low enough latency for real-time monitoring applications, and the authors later note that
they plan to integrate IMERG into the system. Is there a reason why this isn’t already done? For example, some practical or product continuity advantage to using GDAS instead of IMERG?

- Figure 6: Can the authors comment on the fact that the coarse resolution global run appears to do better than the high resolution CA run in this data comparison? I find it surprising, given the presumed topographic sensitivity of SCA.

- The authors note various limitations and potential areas of improvement throughout the paper. I would find it useful for this information to be included in a short section near the end of the paper on “Limitations and Future Work” that could describe ongoing FLDAS development activities. While the future work isn’t a necessary component of this data description paper, it would be valuable for the reader to have this information when considering adopting FLDAS to support research or operations.