CAMELS-AUS v2: updated hydrometeorological timeseries and landscape attributes for an enlarged set of catchments in Australia

Keirnan J. A. Fowler[1], Ziqi Zhang[1], Xue Hou[1,a]

[1]Department of Infrastructure Engineering, University of Melbourne, Parkville, Victoria, Australia
[a]now at: Department of Energy, Environment, and Climate Action, East Melbourne, Victoria, Australia

Correspondence to: Keirnan Fowler (fowler.k@unimelb.edu.au)

Author response to reviewers – consolidated and updated

	Reviewer comment	Author response (original public response in black, further notes (if any) in blue)	Line numbers in manuscript (track-changes version)
Reviewer 1			
1	General appreciation First of all, I wish to congratulate the authors for their lasting efforts to make available quality-controlled hydrological datasets: publishing an update of their initial dataset is a great initiative, and I hope that it will be an example for the authors of other CAMEL sets. I downloaded the files to check that they behave well, and I understand that the authors already corrected minor issues that were found. The changes between the v1 and v2 are well documented, the authors took the initiative to put as supplementary file a commented version of the first paper, which allows a very quick appraisal of what has changed for somebody who would have already read in detail the initial article (please only remove the comment inside section 3.6.3).To summarize my opinion on this paper and the accompanying dataset I would say: bravo!	Thank you very much for the kind words acknowledging the work that has been done already, and the value of updating the dataset. Note sure if the journal allows inclusion of the commented version of the first paper, but if so, we will check it again for errors and correct the one noted at 3.6.3 – thanks for that. Given the reviewer found it useful, we have included the marked up version of the first paper in the public repository (VersionComparisonFile.pdf)	n/a
2	Minor comments 1. You provide to precipitation estimates, one from AGCD, the	Good point. In the updated paper, we will provide more information that is relevant to choosing between the two	L239-269

	Reviewer comment	Author response (original public response in black, further notes (if any) in blue)	Line numbers in manuscript (track-changes version)
	other from SILO. The unexperimented user would have appreciated a recommendation on which to use… and if there are political reasons why you cannot give this advice, please mention it.	products, and we will also provide a recommendation, as suggested. This will also include a recommendation for which PET product to use – this in response to comment #2 from the other reviewer.	
3	Also mention the evolution of AWAP to AGCD in Table 1 (the name of the column has changed).	In the paper text, we will briefly describe the evolution from AWAP to AGCD, as requested – thanks.	L191
4	2. I see no mention of karst, though I imagine some of the catchments may be affected. Are there one or two examples that you could provide, just to warn unexperimented en-users that there may be cases where the water balance will be difficult to close?	Thanks, a helpful suggestion. We will provide a summary of affected catchments based on the best available data from Geoscience Australia.	L208-213
5	3. You provide several variables to appreciate the level of anthropogenic influences. Some readers may be willing to analyze catchments that are either “almost natural” or “almost unregulated by reservoirs”. Could you provide a mention on this, and perhaps suggest a threshold on some of the descriptors you provide (i.e. we would consider catchments with impound_fac less than xxx to be almost unregulated).	OK, we will suggest suitable thresholds based on consideration of the two components of human impact on hydrology that are already supplied as attributes, namely flow regime factors and catchment factors (see Table A4 under the category “Anthropogenic Influences”). The former accounts for the impound_fac suggested by the reviewer. We have implemented this, but it is slightly different to the above —we do not base it on the two factors separately (flow regime; catchment) but rather on the combined index. This adopted approach arose out of a review of literature and current practice—specifically, it follows the example of Stein et al. (2002).	L229-237

	Reviewer comment	Author response (original public response in black, further notes (if any) in blue)	Line numbers in manuscript (track-changes version)
Reviewer 2			
1	This manuscript presents an update of the Australian Camels data set. The Camels data sets have become very useful for the modelling community and I greatly appreciate the efforts by the authors. While I am very positive about this publication, I would like to raise a few issues, which might improve the usefulness of this data set.	Thank you for your detailed suggestions to improve the manuscript and dataset. Much appreciated!	
2	Major issues: It might be valuable to provide different options for precip and pot-evap series. However, this also shifts the need to decide on the most suitable data to the user of this data set. So, while I agree that it is a special feature to offer different data variants (as stated on P2L33), I would also say that this is not without problems. It would be good if the authors could provide some guidance on which variants to use as standard. After all, one of the advantages of the Camels data sets is that the user does not have to take this type of decisions and that the modeling community has one common data set. If now each user chosen some different combination of precip and pot-evap then the results of modelling studies will not be comparable after all. So, I would recommend doing some analyses and then recommending which data variants should be used usually.	This is a good point, and was also raised by reviewer 1 (see their comment #2). We will provide more information that is relevant to choosing between the available products, and we will also provide a recommendation, as suggested. The recommendation will cover the choice of precipitation and PET. There is no need to do further analysis to provide this information and recommendation.	L239-269
3	The same argument can also be made regarding the inclusion of catchments, which have actually	We accept the logic of this comment. In the revised manuscript, we will state our	L215-227

	Reviewer comment	Author response (original public response in black, further notes (if any) in blue)	Line numbers in manuscript (track-changes version)
	been removed for (probably) good reasons in the newer HRS ("let users decide", P5L106). The authors would be in a much better position to make that decision than most users. Again, I am asking for some clear guidance on when these catchments should (not) be included.	opinion that the current selection guidelines appear to be well founded, and on this basis, we will recommend the 467 that meet this criteria (note that these catchments are already clearly marked in the dataset). As stated in the manuscript, the remaining catchments are still important to include for compatibility with other studies (some for backwards compatibility to v1, others for compatibility with the studies named at line 115 of the original manuscript). Reviewer 2 asked for a selection of "near-natural" catchments, which we will select from the 467. I (Keirnan) have been pondering this since submitting the public response, since the heart of the issue is whether the dataset authors should seek to make every possible decision to standardise the use of the dataset; or whether (as in the past) it remains the responsibility of researchers themselves to decide on data inclusion, depending on context. Following these deliberations we have elected to soften the recommendation slightly (relative to the above) so that, following a discussion of the issues, it says *"We recommend that researchers give due consideration to these matters, including the option of using the smaller subset of 467 catchments from HRS2022."* We hope the added text will at least give users the right information to make an informed decision.	

	Reviewer comment	Author response (original public response in black, further notes (if any) in blue)	Line numbers in manuscript (track-changes version)
4	Much of the manuscript is based on McMahon et al. (under review). In the reference list, no information is provided about which journal this manuscript is under review. Honestly, I find it problematic to base a method section on a manuscript that is under review. Until a paper is accepted/available, I would ask the authors to provide more details on the methods and data used for this part of the analysis. Frankly, I was not able to follow/assess this part.	Firstly, we apologise for omitting the journal information for this submitted manuscript (the answer: HSJ), along with the submission date. One of us (Keirnan) has corresponded with the authors of this manuscript and has been advised that it is currently with the journal, having been subjected to very minor revisions in the most recent resubmission. As such, the authors expect that it will be accepted for publication soon. In the unlikely event that it is still not accepted at the time when the resubmission of the CAMELS-AUS v2 manuscript is due, we will confer with the ESSD editor on the best way forward. The paper has been accepted for publication and the citation is now provided in the reference list.	L171-178
5	Minor issues: In the abstract (and elsewhere): Use past tense for your work, "streamflow and climatic information WERE updated" and "information WAS improved"	We will review the tense as requested, but we won't be using past tense because it is difficult to distinguish between different versions of the same dataset (eg. if we say the information "were" updated, does this refer to the current dataset or a previous one?). Instead, we will use "has been" and "have been" (technically this is the present perfect continuous tense).	eg. L10-13
6	The use of a reference in the abstract looks rather unusual	This follows the convention from the first version of the dataset. If I (Keirnan) recall correctly, we were instructed to do this.	L15

	Reviewer comment	Author response (original public response in black, further notes (if any) in blue)	Line numbers in manuscript (track-changes version)
7	P2L36 “this need”, not clear which need this refers to	We will alter to “Responding to the same imperative to create combined datasets…”	L36-37
8	Table 1: I don’t see what the 3rd column (reason/motivation) adds and would recommend to remove this column	On the contrary, we suggest that users of datasets may find it quite difficult and annoying if a change is made without the rationale being clear. Thus, we propose to retain this information.	L65
9	P4L79: good that the authors clearly state that the selection was done by the ABM, but it would be good to know a bit more (criteria, reference)	Although there is no reference per se, there is a website, and we will add the words “for more information, please refer to <website>”. As for the criteria themselves, they are stated in manuscript in the lines that follow.	L92
10	Figure 1: show a white and a grey circle in the legend. Now you show a black circle and only mention the colors in text. I found this confusing.	OK, we will alter this as instructed.	L82
11	P5L103: unclear what “these 43” refers to	We will add the text “these 43 failed catchments” here. Note that, three lines previous, there are the words “Of the 222 HRS-2015 stations … 43 failed the new selection guidelines”.	L114
12	P7L151: is *.mat a suitable format for sharing data?	Firstly, we note the wider context here; namely, the full text is: *“We ran all the calculation functions in TOSSH [which is a MATLAB-based toolkit written by third party authors] and obtained a unique set of 49 streamflow signatures… Among these, 10 signatures have multiple outputs, so we stored only the 39 single-output signatures in the dataset attribute table. For users who need the complete set, we also provided a .mat file that includes all outputs of TOSSH*	L166

	Reviewer comment	Author response (original public response in black, further notes (if any) in blue)	Line numbers in manuscript (track-changes version)
		including the 49 signatures and associated information such as run-time messages." The key point to note here is that we are providing 39 of the 49 signatures in the regular attribute table (which is a text format). This comment relates only to the other 10 which are not easily stored in the attribute table due to their unusual format; plus also there are runtime messages that are useful to retain. While we agree that .mat is not a universal format, the alternative would be to write "Those users that wish to have the full set of 49 will need to re-run the TOSSH toolbox for themselves", which seems a bit silly when we have this information on file as a .mat file. Further to the above, it is useful to make clear that the code in question (TOSSH) is third party code, which means that the format of TOSSH output as a .mat is not something we can easily control.	
13	Figure 3: Please provide some info on the different variables in the figure caption. Finding the variables and information in the tables is rather difficult, so make things easier for your reader (the maps are useful; help your reader understand them!).	OK, we will augment the caption so that the reader does not need to delve into the large tables for this information. Thanks for the suggestion. We opted to improve findability of the information by shading the relevant variables a different colour in the tables. Initially, we intended to was to augment the caption as above, but in practice it was truly a lot of information to include in a caption. Given that there are fifteen variables	L186

	Reviewer comment	Author response (original public response in black, further notes (if any) in blue)	Line numbers in manuscript (track-changes version)
		being separately mapped here, the caption would end up so long that it would be quite difficult to find information for a given variable—which negates the whole point the reviewer is making! Thus, we hope the compromise solution helps the reader in a way that resolves the reviewer's concern.	