"The CAMELS data set: catchment attributes and meteorology for large-sample studies" *by* Nans Addor et al.

Response to Anonymous Referee #1 by Nans Addor et al.

We thank the reviewer for her/his helpful and positive comments. Before we address them, we would like to stress that since our manuscript was submitted, we added a new set of attributes to CAMELS. We extracted geological characteristics from the GLiM and GLHYMPS data sets and produced catchment-scale averages for the 671 catchments. We think that this addition enables a more complete description of the landscape of the CAMELS catchments, and that it will provide useful insights into hydrological processes. These new attributes are now introduced and discussed in Section 8, Figure 7 and Table 6 – see the end of this document.

We are also pleased to report that the catchments attributes are now freely available online: https://dx.doi.org/10.5065/D6G73C3Q

This manuscript describes and presents a dataset of catchment and hydroclimatic attributes for a set of un-impacted (less impacted?) catchments in the continental USA to facilitate large-sample / comparative hydrologic research. The dataset presented in this paper is a significant contribution to large-sample hydrology and worthy of publication. When combined with the time series records provided by Newman et al. (2015), the CAMELS dataset will allow researchers around the world to quickly test a range of hypotheses without spending significant amounts of time (~months) re-collating a similar dataset. Making the dataset freely available online is excellent and is an example for other researchers to follow. Overall the manuscript is very well written and presented and only requires very minor changes as detailed below.

Minor comments

Abstract: It would be good to indicate in the Abstract that the catchments are un- impacted / less impacted by anthropogenic changes. This important point is buried on Page 12, line 9.

Thank you for the suggestion. We now mention in the first sentence of the abstract that the catchments are minimally impacted by human activities:

"We present a new data set of attributes for 671 catchments in the contiguous USA (CONUS) minimally impacted by human activities."

1

1   *We now also stress this on P2, L27-30, and provide a reference to the section of the N15 paper*
2   *explaining how the catchments were selected, and in particular, how the impacts of human*
3   *activities were assessed:*

4   *"All those catchments have 20 years of continuous discharge record from 1990 to 2009 and are*
5   *minimally impacted by human activities (see Section 2.1 in Newman et al., 2015)."*

6   P3, L3: change "data sets used to for their" to "data sets used for their".

7   *Changed.*

8   P4, L15: remove repetition of "in the" before "Great Plains".

9   *Removed.*

10   P5, L25: The seasonality and timing of precipitation and temperature are summarised by sine
11   curves fit to the monthly mean values. The authors need to report on the goodness-of-fit of these
12   sine curves as the resulting single metric is based on the sine curve's, which may, or may not,
13   provide a good fit to the data.

14   *We added the following text to P5, L26-29:*

15   *"Note that sine curves do not necessarily provide a good fit to the annual precipitation cycle, for*
16   *instance in areas experiencing a strong annual cycle and multiple consecutive months with low*
17   *precipitation, such as California (see Berghuijs and Woods, 2015, for a solution to this issue),*
18   *yet they enable a first-order characterization of the dominant climatological features of diverse*
19   *locations, which is useful for studies such as this one."*

20   *Berghuijs, W. R., & Woods, R. A. (2015). A simple framework to quantitatively describe monthly*
21   *precipitation and temperature climatology. International Journal of Climatology, 36(9),*
22   *3161–3174. https://doi.org/10.1002/joc.4544*

23   P6, L2: Please provide an explanation of what + and - values of the seasonality metric (Figure
24   3a) actually mean.

25   *We added the following to Table 2: "positive [negative] values indicate that precipitation peaks*
26   *in summer [winter], values close to 0 indicate uniform precipitation throughout the year"*

27   P6, L4: "ragnes" should be "ranges".

28   *Changed.*

1   P6, L26: "Hydrograph separation is often considered to be desperate" – although Beven (2012)
2   makes a similar statement, it would be helpful to the reader to know the wider context of this
3   statement. Why is hydrograph separation considered desperate?

4   *We rephrased this sentence to avoid confusion:*

5   *"It has to be recognized that the technique used for the separation influences the estimated*
6   *baseflow index (see Beck et al., 2013; Ladson et al., 2013 for recent examples), yet hydrograph*
7   *separation can provide valuable insights into catchment behavior (e.g., Harman et al., 2011) and*
8   *the base flow index has proven to be a useful variable to compare and classify large samples of*
9   *catchments (e.g., Sawicz et al., 2011; Beck et al., 2016)."*

10   P6, L28: change "provides" to "provide".

11   *Changed.*

12   P7, L14: remove repetition of "as" before "low as".

13   *Removed.*

14   P7, L21: change "slope flow" to "slope of the flow".

15   *Changed.*

16   P8, L6 + other locations: change "Mcmillan" to "McMillan".

17   *Changed.*

18   P9, L1: change "consider important" to "consider it important".

19   *Changed.*

20   P10, L1: change "It however" to "However, it".

21   *Changed.*

22   P10, L16: change "P16both" to "P16 both".

23   *Changed.*

24   P12, L9: change "them is classified" to "them are classified".

25   *Changed.*

P12, L31: change "used in large" to "used in a large".

Changed.

P13, L3: change "catchment are common" to "catchments in common".

Changed.

P13, L11: change "It noteworthy" to "It is noteworthy".

Changed.

P13, L13: change "depend the catchments" to "depend on the catchments".

Changed.

P14, L9: change "it will be keep" to "it will keep".

Changed.

Reference to Ladson et al (2013): change "Bronw" to "Brown".

Changed.

Figure 2 Caption: change "5%of daily" to "5% of daily"

Changed.

Figure 5: To make the comparison between 5g and 5h easier to see, I recommend using a single
scale for the two plots. In this way the same colour would mean the same soil depth in both
maps. In the current version dark green means two different soil depths, which is confusing to
the reader.

We agree that using different scales might be confusing at first, yet because the two data sets
cover very different ranges (0 to 1.5m for STATSGO versus 0 to 50m for Pelletier et al., 2016),
using the same scale leads to a significant loss of details in the one of the two maps. Our
intention with these two maps is to provide an overview of the spatial variations over the
CONUS. Figure 5i allows in contrast for a direct and more quantitative comparison of the two
data sets. We hence decided to maintain the scales used for Figures 5g and 5h.

Figure 5 Caption: change "a-h and j-k" to "a-h and j-l"

Changed.

Figure 6 Caption: change "a) to h)" to "a) to j)"

1   *Changed.*
2
3

1 *Geological characteristics – new section and associated table and figure*

2 **8   Geology**

3 **8.1 Data and methods**

4 We used two complementary global sets to characterize the geology of each catchment. The first
5 data set is the Global Lithological Map (GLiM) by Hartmann and Moosdorf (2012). GLiM
6 synthesizes lithological data from 92 regional maps spread across the globe. The spatial
7 resolution is remarkable, as GLiM relies on ~1.2 million polygons to discretize the Earth surface.
8 Three levels of details are available. In this study, we focus on the first level, while the two other
9 levels provide further details that could be processed at a later stage. The first level differentiates
10 between 16 lithological classes (see the list of classes in the legend of Figure 7). We determined
11 the contribution of each lithological classes to the area of each catchment, and recorded the first
12 and second most frequent class within the catchment, as well as the fraction of the catchment
13 they cover. The class "carbonate sedimentary rocks" is particularly relevant from a hydrological
14 perspective (it designates areas likely to host karst systems), we hence also recorded the fraction
15 of each catchment associated with this class. Finally, note that although a 0.5°-gridded version of
16 GLiM is available, we used the more detailed polygon-based version for this study.
17
18 The second data set we used to characterize catchment geology is the GLobal HYdrogeology
19 MaPS (GLHYMPS) of the subsurface permeability and porosity by Gleeson et al. (2014).
20 GLHYMPS is based on GLiM spatial polygons, so its level of spatial details is equally high.
21 Gleeson et al. (2014) principally relied on GLiM lithologic classes to derive quantitative
22 estimates of two key characteristics of the geologic units below soil horizons: porosity and
23 permeability (i.e., the ease of fluid flow through porous rocks and soils). For CAMELS, we
24 produced catchment-averages of these two variables, the contribution of each spatial polygon
25 being weighted by the fraction of catchment it covers. The arithmetic mean was used for
26 porosity, but for permeability, we followed Gleeson et al. (2011) and used the geometric mean
27 instead. The geological attributes are summarized in Table 6.
28
29 A clear advantage of these high-resolution global lithological maps is that they can be used to
30 extract catchment-scale attributes for diverse parts of the globe. Yet, data quality is spatially
31 variable and caveats of the GLiM and GLYHMPS (outlined in the Section 3 of Gleeson et al.,
32 2014) should be kept in mind. In particular, there are unrealistic spatial discontinuities coinciding
33 with jurisdictional boundaries in GLiM maps, which by construction also affect GLYHMPS
34 maps (for instance in the region of North and South Dakota).
35
36 **8.2 Results and discussion**
37
38 The four most frequent dominant geological classes in CAMELS catchments are siliciclastic
39 sedimentary rocks (34% of the catchments), unconsolidated sediments (19%), metamorphics
40 (16%) and carbonate sedimentary rocks (12%). Unconsolidated sediments dominate in
41 catchments along the Gulf Coast and along the southern to middle Atlantic Coast (Figure 7a). In
42 those catchments, both the subsurface porosity (Figure 7f) and permeability (Figure 7g) are high.
43 The Pacific Coast and the region north of the Appalachian Mountains features catchments rich in

siliciclastic sedimentary rocks, leading to a comparatively low subsurface permeability. To the south of the Appalachian Mountains, metamorphic rocks are dominant, resulting in a particularly low subsurface porosity. Finally, the catchments with the highest proportion of carbonate sedimentary rocks are principally located in Central-Western Texas, in the region stretching from Lake Michigan to and including Missouri (Figures 7a and 7f) and to some extent in the Appalachian Mountains (Figure 7b). In addition to these three main regions, there are also isolated catchments with a high proportion of carbonate rocks, for instance in Florida, Nevada and Vermont. The subsurface permeability of those catchments is high. Overall, in 18% of the CAMELS catchments, there is only one GLiM lithological type (Figure 7d), while in 11% of the catchments, the dominant geological class accounts for less than 50% of the catchment area (Figure 7e).

New references

Gleeson, T., Moosdorf, N., Hartmann, J., & van Beek, L. P. H. (2014). A glimpse beneath earth's surface: GLobal HYdrogeology MaPS (GLHYMPS) of permeability and porosity. *Geophysical Research Letters*, *41*, 3891–3898. https://doi.org/10.1002/2014GL059856

Hartmann, J., & Moosdorf, N. (2012). The new global lithological map database GLiM: A representation of rock properties at the Earth surface. *Geochemistry, Geophysics, Geosystems*, *13*(12), 1–37. https://doi.org/10.1029/2012GC004370
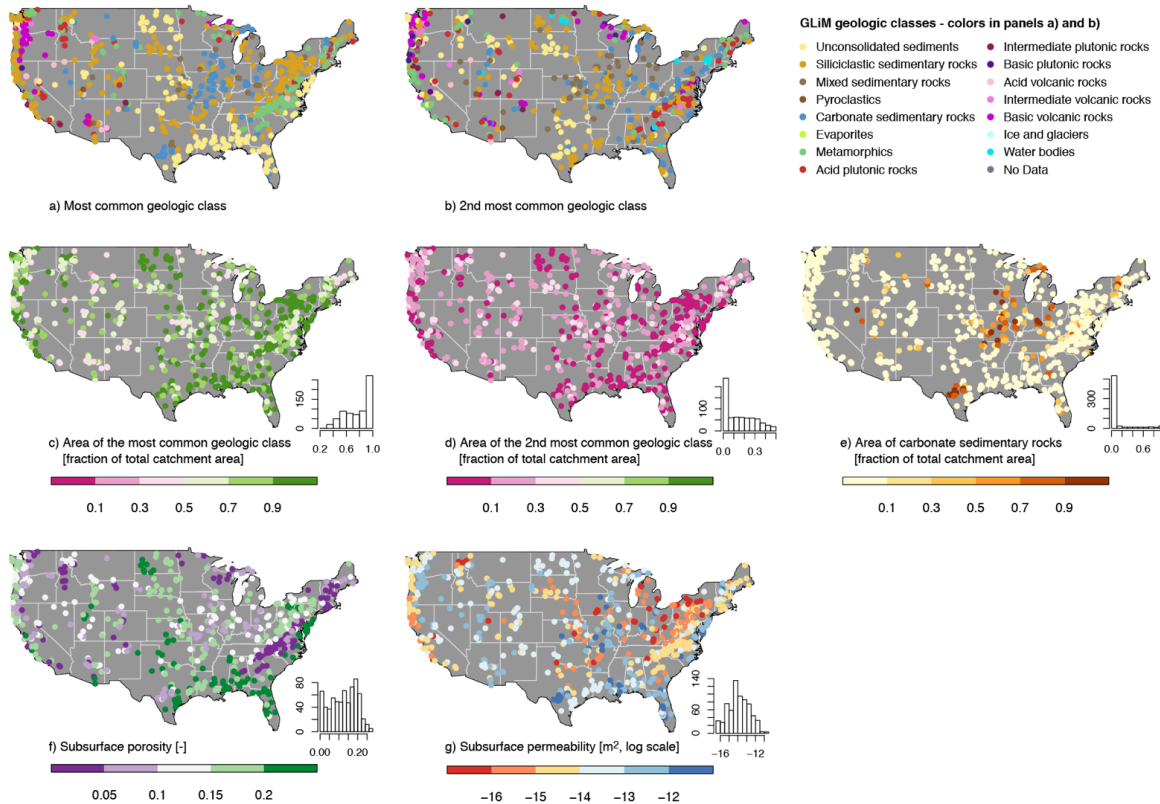
Figure 7: Maps of the geological characteristics over the CONUS. The histograms indicate the number of catchments (out of 671) in each bin.

7

Table 6: Geological characteristics.

Attribute	Description	Unit	Data source	References
glim_1st_class	most common geologic class in the catchment	qualitative	GLiM	Hartmann and Moosdorf (2012)
glim_1st_frac	fraction of the catchment area associated with its most common geologic class	-	GLiM	Hartmann and Moosdorf (2012)
glim_2nd_class	2nd most common geologic class in the catchment	qualitative	GLiM	Hartmann and Moosdorf (2012)
glim_2nd_frac	fraction of the catchment area associated with its 2nd most common geologic class	-	GLiM	Hartmann and Moosdorf (2012)
glim_carb_rocks_frac	fraction of the catchment area characterized as "Carbonate sedimentary rocks"	-	GLiM	Hartmann and Moosdorf (2012)
glhymps_porosity	subsurface porosity	-	GLHYMPS	Gleeson et al. (2014)
glhymps_permeability	subsurface permeability (log10)	m2	GLHYMPS	Gleeson et al. (2014)