We would like to thank the reviewer for the very helpful and constructive comment on the manuscript “Using hydrological and climatic catchment clusters to explore drivers of catchment behavior”.

(comments of the reviewer are printed in blue, responses of the authors are held in black, added text to the manuscript is in italic)

## Response to Reviewer #1 (Anonymous)

Summary

The authors attempt to address the question: “where is hydrologic behavior similar across the contiguous United States?” They use Principal Component Analysis and quadratic regression to cluster catchments in the CAMELS data set located in the Contiguous United States. The variables of interest are 6 hydrologic signatures that earlier research has shown to have high spatial predictability for this dataset. The authors use 15 catchment attributes that were shown to strongly correlate to these 6 signatures to explain the generated clusters in terms of catchment similarity. They discuss which attributes are most influential in determining each cluster and take some steps towards interpreting this from a hydrological processes point of view.

I have read this paper with great interest. I find the separate correlation plots between the east and west CONUS (Figure 6) very interesting and am curious about the differences in eco-regions and hydrologic behavior between the east and west that this might imply.

We are very happy with the interest of the reviewer in our manuscript and like the idea of including the ecoregions in Figure 4. As it turned out, there is a strong correlation and we decided to change the background of the cluster map to the level 1 ecoregions of the Environmental Protection Agency.

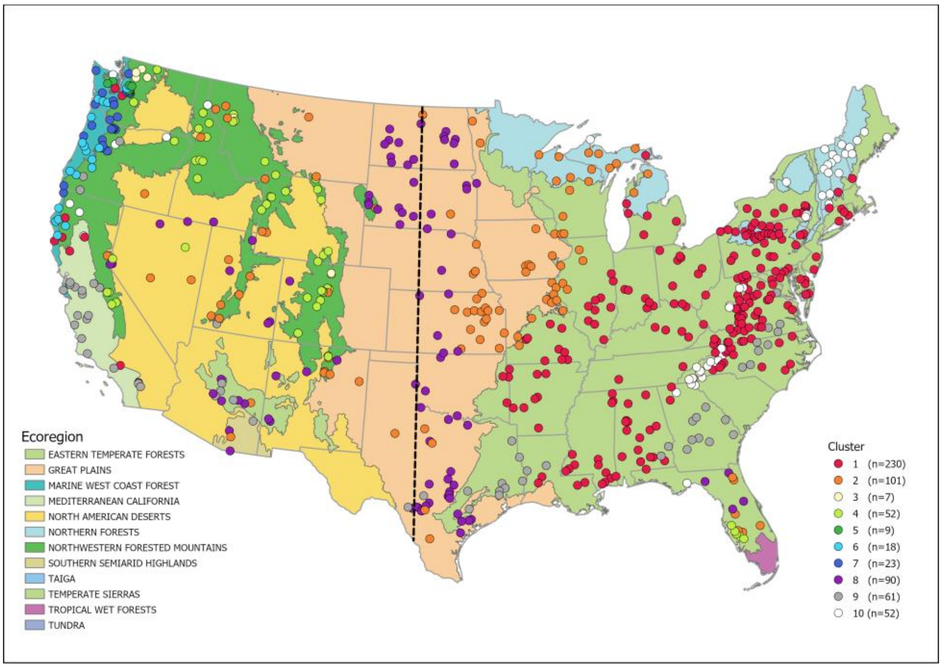



**Figure 4: Locations of the clustered CAMELS catchments and level I ecoregions (Omernik and Griffith, 2014) in the continental US. Dotted line marks the 100$^{th}$ meridian.**

To put a larger emphasis on the correlation plots between the east and west CONUS we changed them from Figure 6 to be Figure 1. Additionally we added a chapter at the beginning of the results and discussion, where we describe the correlations between the catchment attributes and the ecoregions. We also include the precipitation seasonality as an additional catchment attribute in this figure. This section reads as follows:

*3.1 Catchment attribute correlations in the CAMELS data set*

*Usually the 100th meridian is seen as the dividing climatic line in the US, splitting the country in a semi-arid west and a humid east. We assume that this difference in climate also has implications for the hydrology and the overall catchment attributes in those regions. To quantify this we split the CAMELS data set into a western and an eastern part, based on the 100th meridian (Figure 1 and 4). This shows that many of the catchment attribute correlations do not differ much between the east and the west. In most cases (>80%), Pearson correlation coefficients vary by less than 0.4 (Figure 1c). Still, there are some catchment attributes with larger differences of up to 0.7 between both regions. Most striking are the mean elevation and the fraction of the precipitation falling as snow as well as the vegetation attributes LAI maximum and Green vegetation fraction maximum. Even though these attributes are directly related to each other through temperature gradients, they differ substantially in both parts of the country. In the mountainous western US, elevation is highly correlated with the fraction of precipitation falling as snow (r=0.8), while it is not in the eastern US (r=0.4). This, and the different correlations between vegetation and elevation are probably caused by the fact that the temperature gradients differ in both regions. In the western US it is much more mountainous and thus temperatures typically change with elevation. In the more level eastern US, on the other hand, the change in temperature is mainly linked to the latitude. Striking are also the changes of correlation with regard to the fraction of precipitation falling as snow. Here we find altered directions of the correlation, i.e., positive correlations with LAI maximum and frequency of high precipitation events in the east turn to negative ones in the west. It also becomes obvious that all three measures of vegetation seem to track similar characteristics in the catchments, as they highly correlate with each other (especially in the eastern US with r=0.9). In addition, all vegetation attributes depict a large negative correlation with aridity. Hence, the vegetation attributes considered are likely good proxies for aridity. Overall, we see that the relations between the catchment attributes are quite similar for the eastern and western US, with the exception of the mean elevation, snow and the LAI maximum.*

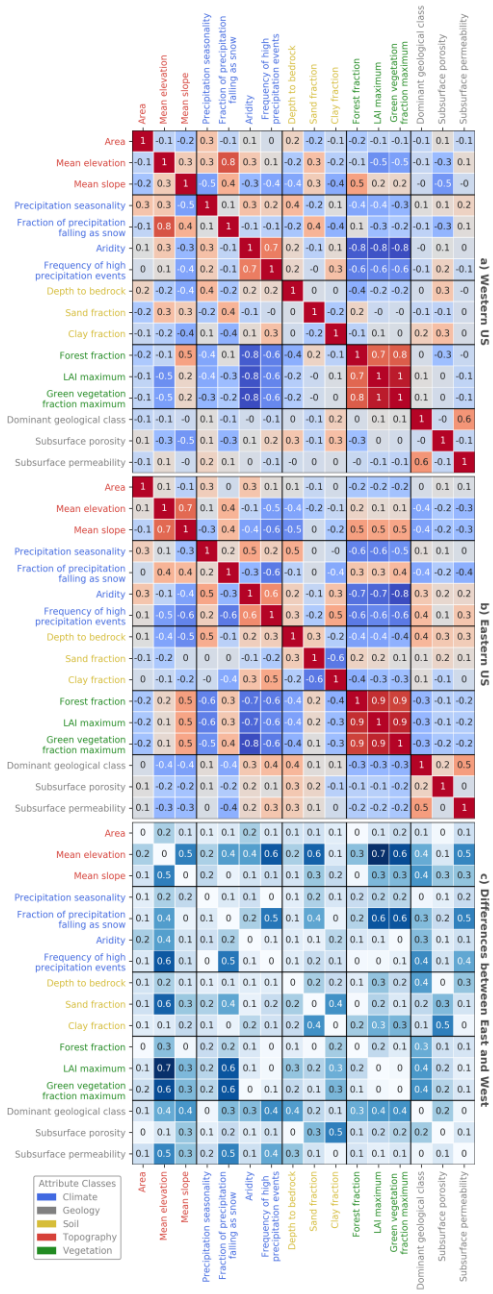



**Figure 1: Pearson correlation coefficients given for all catchment attributes in western (a) and eastern (b) US. Absolute differences of the correlation coefficients between the eastern and western US is given in c). Eastern and western is defined by the 100th meridian. Due to rounding effects, correlations with the same Pearson correlation coefficient might show slightly varying color codes.**

My main concern is that apart from Figure 6, the manuscript mostly seems to confirm earlier work by e.g. Addor et al. (2018), Berghuijs et al. (2014), Knoben et al. (2018) and Kuentz et al. (2017). Confirming findings is not a bad thing, but I think the authors are missing out on an opportunity to go beyond these studies. The authors spend some time in the main manuscript (L219-229; L256-258; Table 2 to some extent) speculating about hydrologic behavior in each of their clusters. More of these thoughts are hidden in the appendices (L460-511). I believe the manuscript would become much stronger if the authors would make this the main topic of the manuscript and spend more time on trying to understand the hydrologic behavior each cluster represents in terms of dominant processes, as this would be a novel contribution to the field. This could be structured similar to

Berghuijs et al. (2014) but the CAMELS dataset gives the authors the catchment information needed to go beyond that work. Addor et al. (2018) could also help to outline potential changes to the manuscript.

The reviewer is right here. An in depth analysis of the processes within the clusters was missing so far. In the revised version, we added a new chapter where we discuss the clusters with their processes in depth:

3.4 Location and properties of the catchment clusters

*The catchment attributes in the CAMELS and similar large scale datasets often show a pattern that resembles climatic zones (Addor et al., 2018; Coopersmith et al., 2012; Yaeger et al., 2012). For the catchments clusters presented here, we can see that most of the clusters roughly follow ecoregions in the US (Figure 4). Especially clusters 1, 4, 6 and 7 are almost entirely located within one ecoregion. Cluster 2, 8 and 9 on the other hand follow those ecological boundaries to a lesser degree.*

*We can see a split of the clusters along the 100th meridian. Cluster 3, 4, 5, 6 and 7 are located mainly in the west, while Cluster 1 and 10 are mainly found in the east. However, the remaining Clusters 2, 8 and 9 have roughly similar numbers of catchments in both regions. Overall, catchments in the eastern half of the United States form large spatial patterns of similar behavior, while the catchments in the west are patchier.*

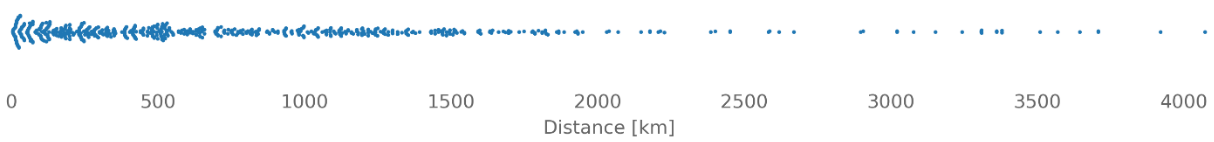



**Figure 5: Swarm plot of the real world distances of all catchments to the most hydrologically similar catchment (based on their distance in the PCA space of the hydrological signatures).**

*In addition, similar catchments can be quite far away from each other (Figure 5). Sometimes, the catchment with the most similar signature was found as far as 4,000 km away (almost the entire longitudinal distance of the continental US). This explains why spatial proximity seems to be important in some studies that look into explanations of catchment behavior (Andréassian et al., 2012; Sawicz et al., 2011), but not in others (Trancoso et al., 2017). This also indicates that clustering by using spatial proximity might only work in regions like the eastern US, where the behavior of rivers only changes gradually. The finding that the most similar catchment (based on their hydrological signatures) can be far away, also explains the behavior of clusters that contain catchments quite distant from each other (e.g. Cluster 4). Even though the catchments might be far away from each other, the interplay of different catchment attributes and driving factors, including sometimes very different climates, can lead to similar (equifinal) discharge behavior concerning the overall amount of discharge, its distribution in the year, the high flows and the runoff-ratio. This was also found by several other studies (e.g. Berghuijs et al. (2014); Knoben et al., 2018; Kuentz et al., 2017)).*

*In the following, we describe the catchment clusters in regard to their characteristics in meteorology (Figure 6), attributes (Figure 7), hydrology (Figure 8) and location (Figure 4). The main points of this description are summarized in Table 2. A list of all catchments with index, position, cluster classification and climate indexes is given in the supplementary material.*

*Cluster 1 is defined by a dense vegetation cover (Figure 7). The low elevation of those catchments results in little annual snow fall. They are mainly located in the southeastern and central plains and therefore get relative high rainfall (>1,000 mm per year) (Figure 4), almost uniformly distributed over the year (Figure 6). From a hydrological perspective, these catchments produce little discharge. Cluster 1 contains the highest number of catchments (n=230). So over one third of the catchments in CAMELS show a relatively similar behavior when it comes to the amount of water fluxes and their distribution throughout the year. This is particular visible when we look at annual supply of discharge (Figure 6). Even though the cluster contains a large number of catchments that also partly differ a lot in their potential evapotranspiration, there is only a minor difference in the amount of discharge and its seasonality.*

*Cluster 2's most typical attribute is its high precipitation seasonality. However, concerning most other catchment attributes, Cluster 2 is undefined as it contains catchments of most regions of the continental US (with a concentration in the eastern Great Plains) (Figure 4). The hydrological signatures on the other hand show a clearer pattern. Here, the mean winter discharge, Q95 and the mean annual discharge have a narrow range (Figure 8). This shows that catchments with very different attributes can produce similar discharge characteristics. The different attributes seem to cancel each other out in their influence on the discharge. This might be enhanced by the high precipitation seasonality with higher precipitation in the summer, which creates a strong climatic forcing and thus a narrow range for the hydrological signatures (Figure 6). This cluster differs from the first one, by having even lower discharge, with almost no peaks and a higher influence of snow melt.*

*Cluster 3 is the smallest cluster with only seven catchments. Those are all located in the Northwestern Forested Mountains. Their most distinct feature is their strong negative precipitation seasonality (indicating a strong precipitation peak in the winter) (Figure 6, 7). They also experience high precipitation events mostly in winter falling as snow. Hydrologically, their most distinct features is the very high mean summer discharge and high runoff ratio (Figure 8). This is probably caused by the large amounts of snow melt in late spring and early summer. The catchments of Cluster 3 have the largest overall snow storage with mean maximum value of over 600 mm. Overall, the catchments in this cluster seem to be, from a hydrological point of view, the most extreme in the overall CAMELS data set. This can be seen in their varying discharge patterns. The uniting pattern is their large peak discharge during summer and their extreme values in the PCA space (indicating much higher values for the hydrological signatures in comparison with the other catchments) (Figure 3).*

*Cluster 4 is, as cluster 3, located in the Northwestern Forested Mountains, with the exception of four catchments that are located in Florida (Figure 4). This cluster is another example of different catchment attributes being able to create similar discharge characteristics concerning the signatures used, while having very different catchment attributes (Figure 6). The catchments have overall low discharge and few high flow events, except one large peak in the mid of the summer, which is caused by melting snow in the northern catchments and strong rainfalls in Florida. Their catchment attributes vary widely, especially in all attributes that are related to elevation (e.g. fraction of precipitation falling as snow) (Figure 7), which is to be expected when some of the catchments are located close to the sea in the southeast, while others are mountainous.*

*Cluster 5 includes only few catchments (n=9). They are all located at regions in the northern part of the Marine West Coast Forests (Figure 4). This is the region in the continental US that receives the highest precipitation (>2000 mm year), which is reflected in their discharge characteristics (Figure 6, 8). These catchments have the highest discharge in the whole dataset, especially in the early summer, due to a combination of high precipitation and snowmelt. They also experience only few high* *precipitation events as they receive large amounts of rain and snow most of the year, with a distinct very high peak in the winter months. The catchments are uniformly covered by almost 100% of forest.*

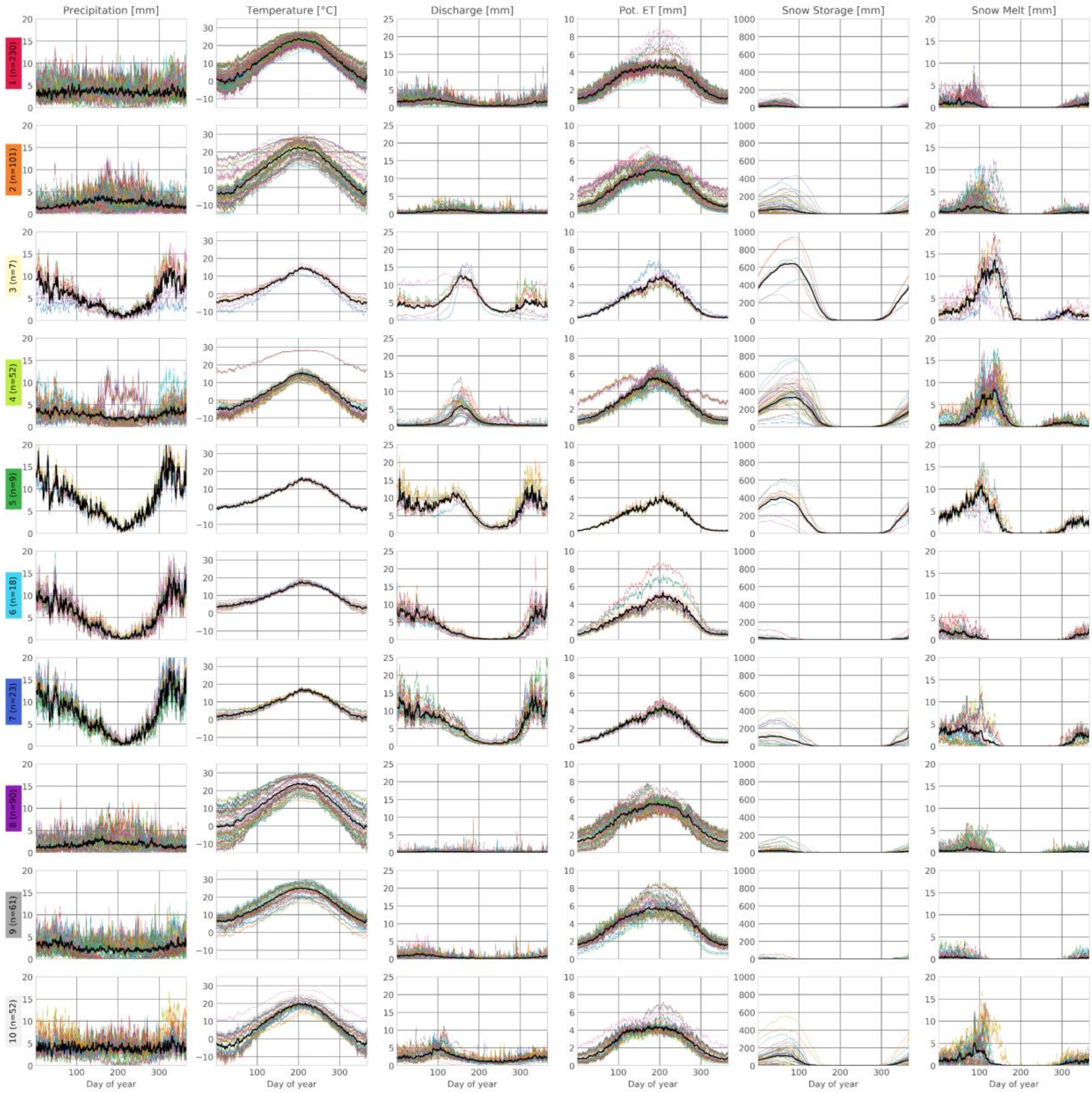



Figure 6: Meteorological attributes of the clustered CAMELS catchments averaged by day of the year. Potential Evapotranspiration (Pot. ET) was calculated with Hargreaves-Samani (Samani, 2000). Snow storage and melting was calculated using a temperature based approach described (Massmann, 2019). Black lines indicate the mean of all cluster members. Colored lines represent the individual catchments.

*Cluster 6 is located in the Marine West Coast Forest, but in contrast to Cluster 5, they cover the whole region and not only the northern part (Figure 4). The catchments are very similar in their attributes and discharge characteristics to Cluster 5, with the exception of lower discharges and runoff ratios (Figure 7, 8). This is caused by slightly lower precipitation in comparison with Cluster 5. Cluster 6 experiences the most negative precipitation seasonality across all clusters, with almost all precipitation falling in the winter month. Due to this seasonality and the lower precipitation in the summer, the catchments of this cluster uniformly dry out almost completely in the late summer (Figure 6).*

*Cluster 7 is also located in the same region as Cluster 5 and 6 (Marine West Coast Forests) (Figure 4). Concerning the catchment attributes and the discharge characteristics, it is located between Cluster 5 and 6. So, Cluster 5 to 7 all cover the same region and differ in their mean summer discharge, which is caused by variations in elevation and location (Figure 7). Cluster 7 has higher subsurface permeabilities than cluster 6, which might explain the differences in hydrological behavior, even though the overall attributes of both clusters are rather similar. For example, Cluster 7 has an overall lower discharge than Cluster 5, but does not dry out during the summer as Cluster 6 does (Figure 6). This might be due to the larger amount of snow it receives in comparison with Cluster 6 and its lower evapotranspiration.*

*Cluster 8 is the most arid cluster (Figure 7). All of the catchments are located in western parts of the Great Plains and in the North American Deserts (Figure 4). They are characterized by an overall low water availability and high evaporation, which is shown in the very low mean annual discharge and runoff ratio (Figure 6, 8). This also results in low values for the LAI. However, the frequency of high precipitation events is high.*

*Cluster 9 covers all southern states of the United States (Figure 4). The catchments here are quite similar to Cluster 8, but show a lower precipitation seasonality and a higher forest cover and green vegetation (Figure 7). In addition, all catchments of this cluster are in relative close proximity to the sea. The uniting factor in this cluster seems to be the very low snow fraction and the high evapotranspiration (Figure 6, 7).*

*Cluster 10 catchments are all located in the Appalachian Mountains (Figure 4). The mean elevation is higher than of most other clusters and the catchments also depict a low aridity and a very high forest cover (Figure 7). Their discharge characteristics are similar to that of the Marine West Coast Forests Clusters 5 to 7 (Figure 6, 8). However, they receive less water than those catchments. Cluster 10 covers the same ecoregion as Cluster 1, but has a distinct behavior due to its mountainous character, which can be seen in the higher seasonality of the discharge. This is probably caused by the larger snow cover, with a returning snow melt discharge peak in spring.*

*Overall, we can see similar trends for some of the cluster. We identified four distinct groups. The general similarities of the clusters are also represented by their distance and position in the PCA space (Figure 3).*

- *Group 1 (Cluster 1, 2, 8, 9): low seasonality in precipitation and discharge; located in the eastern US; due to low slope inclinations, water takes a long time to reach the outlet.*
- *Group 2 (Cluster 3, 4): dominant summer peak of discharge caused by rapid snow melt; mostly located in the mountains of the western US; differ in precipitation inputs.*
- *Group 3 (Cluster 5, 6, 7): located in the Northwestern Forested Mountains; characterized by high precipitation amount and seasonality, but more or less extreme versions.*
- *Group 4 (Cluster 10): located in the Appalachian mountains; share characteristics with Group 1, though influenced by higher elevations and steeper slopes.*

*Those groups of clusters are similar to the ones identified by (Berghuijs et al., 2014), even though they used a very different method to derive them (based on seasonal water balance and hydro climate). The main difference in the groups is probably caused by how we structure the clusters and groups in the eastern US, due our clusters being more influenced by the Appalachian Mountains. However, both approaches deliver similar results overall.*

*The question remains: what is the right numbers of clusters? Though even we did find four distinct groups, having only four clusters would probably be too little, as the clusters in the groups show a wide range of behavior (Figure 3, 7, 8, Table 2). There are catchment attributes, which we did not take into account, but which could further split up the clusters (e.g. the shape of the catchments). However, this study considered the catchment attributes that are usually considered as being important. The fact that the clusters contain different numbers of catchments can be explained by their distances in the PCA space (Figure 3). Many of the catchments are rather similar. This produces some clusters with most of the catchments. However, we also have some extreme catchments (e.g. Cluster 3 and 5), which are very different to the bulk of the catchments in the CAMELS dataset. Thus, even though some of our presented clusters are quiet small in number, they are needed to capture their extreme hydrological behavior. Our results show that some of the clusters follow the boundaries of the ecoregions in the US very directly (Cluster 1), while others do not (Cluster 9). The worlds of ecology and hydrology are sometimes shaped by the same forcing, but not always.*

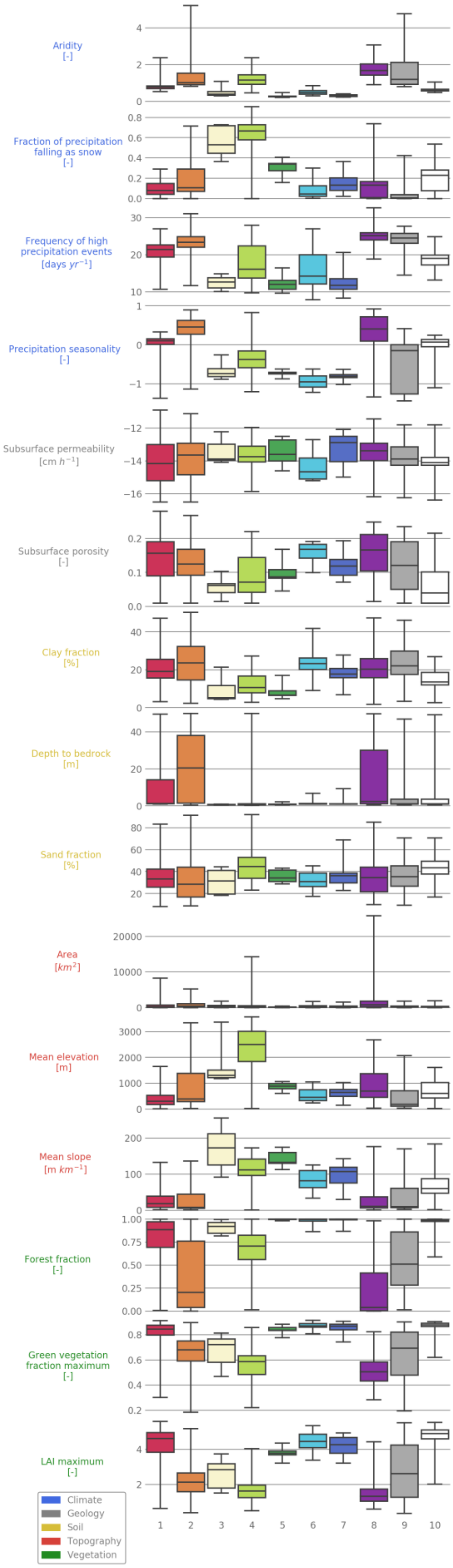



**Figure 7: Boxplots of the catchment attributes of the clusters**

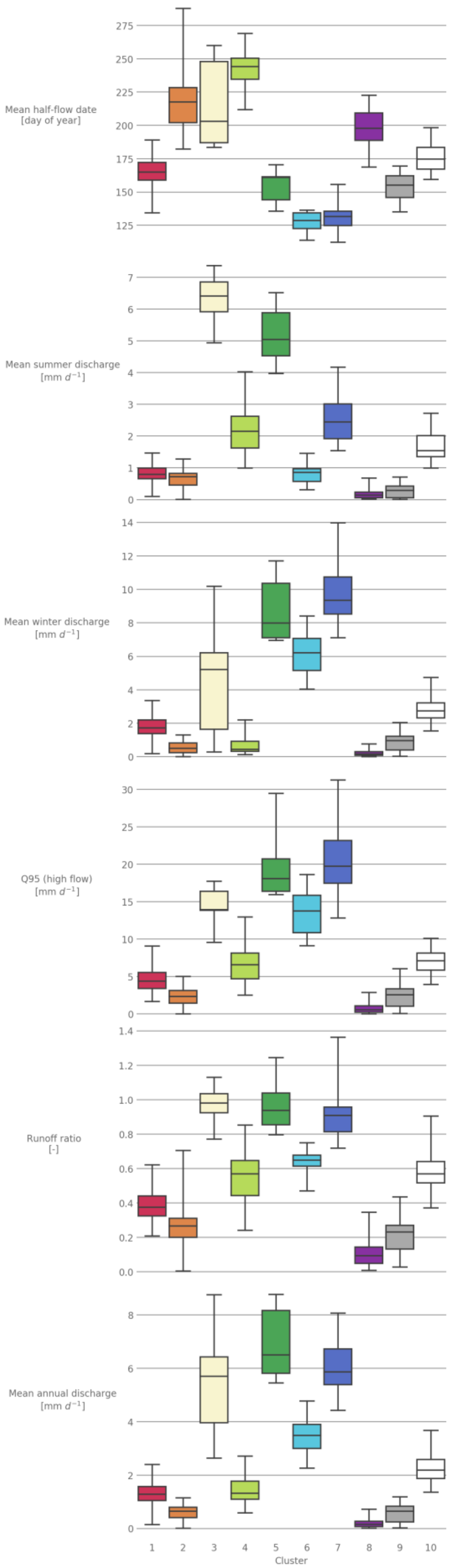



**Figure 8: Boxplots of the hydrological signatures of the clusters**

Table 1: Properties of the catchment clusters. Typical signatures/attributes refers to the signature/attribute of the cluster with the lower coefficient of variation scaled by the mean coefficient of variation of the whole dataset. Dominating attribute refers to the catchment attribute that has the highest weighted $R^2$.

| Cluster | n | Main Region | Typical signature | Typical attribute and their manifestation | Dominating attribute |
|---|---|---|---|---|---|
| 1 | 230 | Southeastern and Central Plains | Low mean winter discharge | Low aridity | Aridity |
| 2 | 101 | Central Plains (with scattered catchments all over western US) | High mean half-flow date | High precipitation seasonality | Green vegetation fraction maximum |
| 3 | 7 | Northwestern Forested Mountains | High mean summer discharge | Low precipitation seasonality | Fraction of precipitation falling as snow |
| 4 | 52 | Northwestern Forested Mountains and Florida | High mean half-flow date | Mid frequency of high precipitation events | Precipitation seasonality |
| 5 | 9 | Northern Marine West Coast Forests | High mean summer discharge | Very high forest fraction | Forest fraction |
| 6 | 18 | Marine West Coast Forests | Mid runoff ratio | Low precipitation seasonality | Aridity |
| 7 | 23 | Western Cordillera (Part of Marine West Coast Forests) | High mean winter discharge | Low precipitation seasonality | Fraction of precipitation falling as snow |
| 8 | 90 | Great Plains and North American Deserts | Mid mean half-flow date | High frequency of high precipitation events | Precipitation Seasonality |
| 9 | 61 | All southernmost states of the US | Low mean half-flow date | High frequency of high precipitation events | Aridity |
| 10 | 52 | Appalachian Mountains | Low mean winter discharge | High forest fraction | Mean elevation |

I also have a few methodological concerns about the way the signatures and attributes have been selected and how the number of clusters has been determined, and how these choices might limit the authors' ability to go beyond these earlier studies (see below).

Major comments

1. The authors use the six most predictable signatures from Addor et al. (2018) for analysis here. They use 15 catchment attributes that Addor et al. (2018) indicates as having the highest ability to explain these signatures, with climatic attributes having the strongest connection to signatures. Earlier work (Berghuijs et al., 2014; Kuentz et al., 2017; Knoben et al., 2018) has also shown the strong influence climatic conditions have on hydrologic behavior and advocate for further studies

that investigate the impact of less clear relations between catchment attributes, such as resulting from geology or vegetation, and hydrologic behavior. These relations can be seen in well-monitored experimental catchments and must logically exist in all other catchments, but in large-sample studies this has so far not been conclusively shown (Addor et al., 2018, provides various compelling reasons for why that might be the case). Progress towards understanding these relations would be an important contribution to the literature and interpreting cluster analysis from a dominant hydrological process perspective could be a first step. The authors take some steps in this direction, but unfortunately they are limited by their study setup to mostly confirm what has already been shown before, without having the necessary information available to go beyond these earlier studies. Restructuring of some of the study setup and analysis might be needed (see points 2 and 3 below).

We agree and have included a more complete description and interpretation of our clusters in section 3.4, see previous comment. In addition, we have rewritten section 3.5:

*3.5 Importance of the catchment attributes in the clusters*

*The individual importance of the catchment attributes in the clusters is variable and partly deviates from the order of importance in the overall dataset (compare Figure 2 and Figure 9). For Cluster 1 (Southeastern and Central Plains), 6 (Marine West Coast Forests) and 9 (coastal states) aridity has the highest weighted coefficient of determination in the clusters. For Cluster 3 (Northwestern Forested Mountains) and 7 (Western Cordillera) the highest relevance is found for the fraction of precipitation falling as snow. For the remaining clusters it is precipitation seasonality (Cluster 4 (Northwestern Forested Mountains), Cluster 8 (Great Plains and Deserts)), the green vegetation fraction maximum (Cluster 2 (Central Plains)) and the mean elevation (Cluster 10 (Appalachian Mountains)). We can also see that some clusters have one dominating catchment attribute (investigated by the coefficient of determination e.g. aridity in Cluster 1, compare Figure 9), while for other clusters, all attributes seem equally important (e.g. Cluster 8). Overall, the western clusters (west of the 100th meridian) display the highest weighted* $R^2$ *with:*

- *Fraction of precipitation falling as snow (Cluster 3, 7)*
- *Precipitation seasonality (Cluster 4)*
- *Forest fraction (Cluster 5)*
- *Aridity (Cluster 6)*

*eastern clusters (east of the 100th meridian) with:*

- *Aridity (Cluster 1)*
- *Mean elevation (Cluster 10)*

*clusters equally present in west and east with:*

- *Green vegetation fraction maximum (Cluster 2)*
- *Aridity (Cluster 9)*
- *Precipitation seasonality (Cluster 8)*

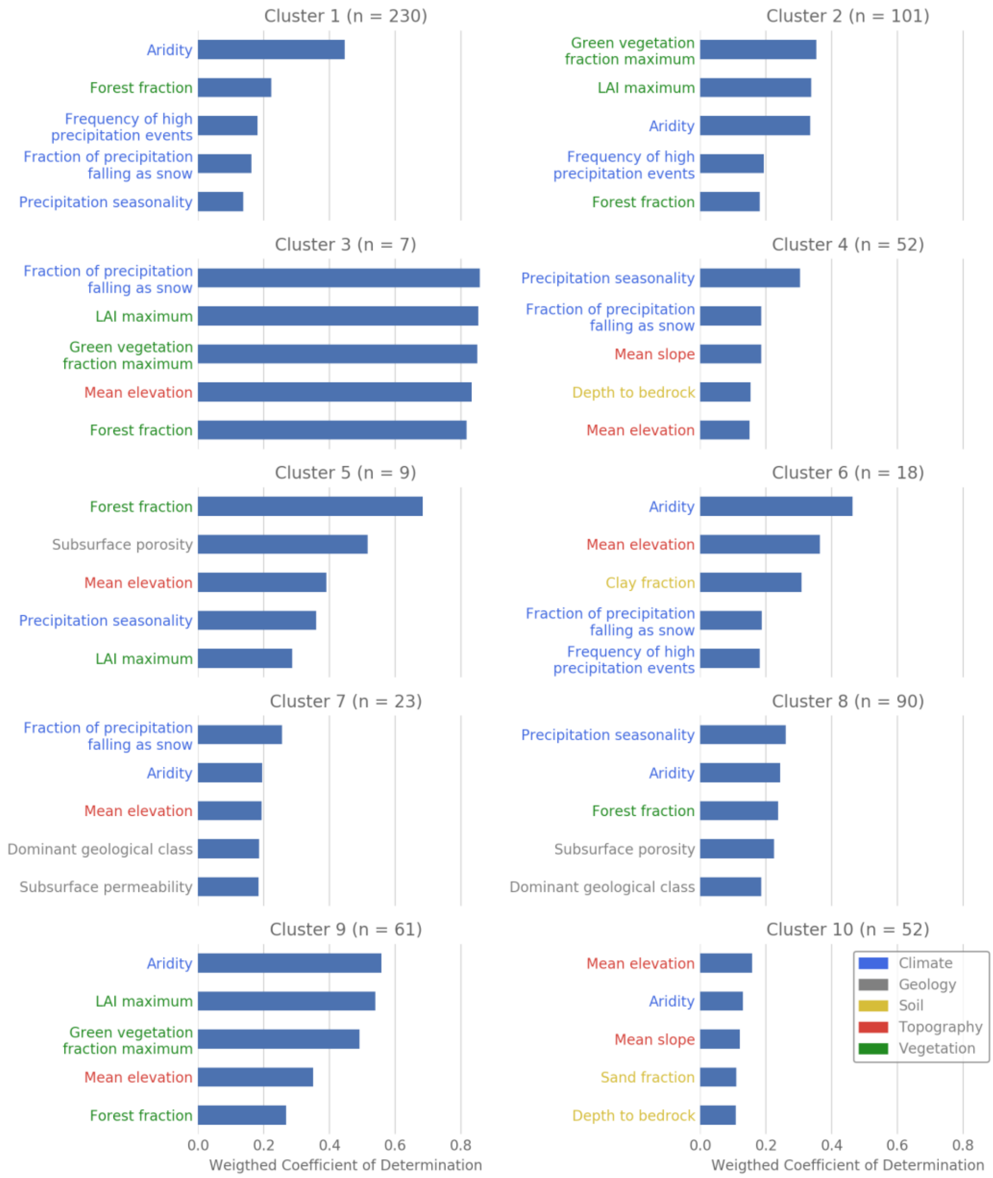



Figure 9: Importance of the catchment attributes evaluated by the quadratic regression for the catchment clusters. Attributes colored according to their catchment attribute class.

*Keeping the correlation coefficients displayed in Figure 1 in mind, we see that climate is the most important factor in almost all clusters, as the vegetation attributes are highly correlated with the climate attributes. The only exception is Cluster 10 in which mean elevation is the most important catchment attribute. However, the catchment attributes in Cluster 10 have overall low $R^2$ values and the mean elevation is directly followed by the aridity. This again shows that climate seems to be the dominating factor for catchment behavior, as found in other large sample studies (e.g. (Berghuijs et al., 2014; Kuentz et al., 2017)). Nevertheless, if one takes a closer look at the data set, more detailed, regional correlations with regard to individual climate variables can be determined. For example, Cluster 1 is defined by the aridity, while Cluster 4 seems to be much more influenced by the precipitation seasonality. Overall, it is feasible to link dominating catchment attributes to the hydrological behavior. While it is straightforward in some regions of the US, it is more challenging in others. We link this to the signal of the climatic forcing being more superimposed by other catchment attributes, which results in a less clear connection between its hydrological behavior and the climate. This hints that climate and catchment attributes are more intertwined in those areas and indicates regions where different types of hydrological runoff generation processes are existing. Furthermore, it* *indicates regions where hydrological predictions in ungauged basins (Hrachowitz et al., 2013) can become very challenging, as the interplay of the available meteorological data and catchment attributes cannot sufficiently explain the hydrological characteristics. Those findings also highlight one current discrepancy between large sample and single catchment studies. While large sample studies, especially the very large ones, identify climate as being most important for the hydrological behavior (e.g. (Addor et al., 2018; Kuentz et al., 2017)), smaller sample studies (e.g. (Chiverton et al., 2015; Pfister et al., 2017)) and single catchment studies (e.g. (Floriancic et al., 2018) often identifiy the geology or soils as being very important. This might be linked to the overall problem of scales in hydrology as different scales of soil/geology and climate have different effects and different data accuracy (Addor et al. 2018). In addition to this, the overall scale might also come into play. Smaller studies often compare catchments that are not far away from each other and probably have similar climate forcings. Thus, the differences in hydrological behavior can only be caused by catchment attributes other than climate. Therefore, larger and smaller sample studies might be looking at different things. While very large sample studies capture what drives catchments on large scales (the climate), smaller studies look at how this climatic signal is transferred to discharge by the catchment attributes.*

2. Hydrologic interpretation is limited by the choice of signatures. I appreciate that the authors were looking for signatures that are easily predictable in space, but this limits the generality of the conclusions that can be drawn. The chosen six signatures do not describe the full hydrologic regime, focusing mostly on flow magnitude (mean annual, summer and winter flow, runoff ratio, Q95) and somewhat on seasonality (half-flow date), with no signatures dedicated to low flows, intermittency of flows, or response time of the catchment. Therefore statements such as the following are too general for the supporting analysis and should be rephrased to account for the specific conditions these 6 signatures describe (please note that this list could be incomplete):
- L95. “These two principal … overall hydrological behavior.”
- L188. “ … lead to similar (equifinal) discharge behavior”
- L464. “So over one third of the catchments in CAMELS show a relatively similar behavior.”
- L470. “… catchments with very different attributes can produce very similar discharge characteristics, …”
- L480. “an example of different catchment attributes being able to create similar discharge characteristics concerning their signatures, while having different catchment attributes”

We changed those statements to make clearer that they refer to the signatures used in this study.

Related to this, both Addor et al. (2018) and Knoben et al. (2018) show that these particular signatures correlate strongly with climatic conditions in the catchment. I doubt whether there is much to be learned about the influence of non-climatic attributes on hydrologic behavior by looking only at signatures with such strong connections to the prevailing climate. Using a wider range of signatures could allow more in-depth analysis of the relation between attributes and signatures. E.g. McMillan et al. (2017) could be of use in choosing different signatures:

McMillan, H., Westerberg, I., & Branger, F. (2017). Five guidelines for selecting hydrological signatures. Hydrological Processes, 31(26), 4757–4761. https://doi.org/10.1002/hyp.11300

We agree that a different set of hydrological signatures might lead to different sets of clusters. However, we did focus on the signatures, as we wanted to know if the signatures identified by Addor et al. (2018) can be used to create meaningful hydrological clusters, which is part of the main research question in this manuscript. Changing the signatures now would lead to an entirely different paper. Consequently, we removed statements in the manuscript, which might have implied a too general picture of the captured hydrological behavior.

Despite not re-clustering with additional signatures, we picked up the idea of investigating the clusters performance with regard to low flows. The new figure 6 provides information how the different discharge patterns are captured by our clusters. Interestingly, it turns out that the catchment's behavior during low flow conditions is very similar in the individual clusters, although we have not included signatures that are concerned with low flows. We conclude that the hydrological signatures we used contain already sufficient information to present in the discharge patterns.

3. Hydrologic interpretation is also limited by the choice of attributes, because the selected attributes are strongly correlated with one another. I had already written a few comments on this before reaching Figure 6, which shows that the authors are aware of these correlations. This knowledge should play a much larger role in the earlier parts of the paper, where the study setup is decided (i.e. which attributes to use) and where the importance of attributes for clustering is discussed (for example, the 5 most important attributes in cluster 3 are essentially 2 factors spread out over 5 attributes: snow & elevation are the first (r = 0.8), and various aspect of vegetation are the second group (r=0.7, r=1 and r =0.8). A different selection of attributes might be needed. I also believe that enforcing 3 attributes per attribute category is unnecessarily limiting and ignores some of the current understanding of drivers of hydrologic behavior, such as not using a climate seasonality metric (further details below).

In line with comments and replies above, we now focus on the former figure 6 (which is now figure 1) in this manuscript. A new in-depth analysis on the correlations in the data set is given (section 3.1). Further, we include the precipitation seasonality as a climate seasonality metric that is provided in the CAMELS dataset.

Further missing catchment attributes, which are present in the CAMELS dataset, such as the fraction of carbonate rocks or the water fraction in the soil, have not been mentioned in any of the literature we have reviewed as being very important for hydrological behavior.

Minor comments

L45. Addor et al. (2018) identified these signatures as having low spatial predictability in the US. Is it correct to assume that these conclusions also apply to the study domain of Kuentz et al. (2017), i.e. Europe?

To be honest, we do not know. Obviously, Europe and the continental US are different in many aspects, but we would be surprised if they are so different that hydrological findings cannot be transferred between those regions, especially when they are derived from large data sets.

L68. Is there a reason to assume that the most diverse total information is retained by using 3 attributes each from climate, topography, vegetation, soil and geology? To what extent are all

CAMELS attributes correlated and to what extent is the subset of 15 attributes correlated with one another?

This point is captured in our answer to the main point 3 of the reviewer.

L69. [Adding to the previous comment] Among others, Berghuijs et al. (2014), Addor et al. (2018) and Knoben et al. (2018) have found that climatic seasonality is an important control on hydrologic behavior. The authors have included 'frequency of high precipitation events' over a climate seasonality metric. I agree that there can be good reasons to include the frequency of high P events metric but because the authors limit themselves to 3 attributes per attribute category, they cannot include a seasonality attribute even though current theory indicates that seasonality can be an important control on hydrologic behavior. Given this, I think the choice of 15 attributes and how they are distributed between the different categories needs to be better justified and possibly changed.

As suggested by the reviewer, we included precipitation seasonality in our analysis and the revised version of the manuscript. See further replies to this issue in the main points raised.

L95-97. I don't think the two PCA's of the six signatures can be seen as "describers of the overall hydrologic behavior". This sentence and the next one need to be more nuanced, because the authors state in section 2.1 that no low flow signatures are part of their selection. Other possibly relevant aspects of the flow regime, such as baseflow or flashiness, are also not covered in this selection of signatures.

Changed as proposed.

L119. Kuentz et al. (2017) use 10 clusters to group >35000 catchments using 16 different signatures. I expect that choosing to use 10 clusters in this study with >600 catchments and 6 signatures might provide unnecessary granularity. Can the authors somehow quantify the difference between each pair of clusters to show that 10 is an appropriate number? If such quantification is not possible, did the authors investigate the impact of using fewer or more clusters?

[additional note] Seeing that cluster 3 only contains 7 catchments and that cluster 5 only has 9, but that cluster 1 has 230 catchments in it, I think that some more discussion of the number of clusters is warranted. Cluster 5, 6 and 7 also look very similar, possibly indicating that too many clusters have been used. Some questions that come to mind:

- What is the explanatory power of a cluster with only a handful of catchments in it?

We also tried to use the elbow method (that was used by Kuentz et al (2017)) to find the right amount of clusters. However, this did not produce a clear cut answer on how many clusters should be used in our study. Therefore, we assumed that the larger database of Kuentz et al (2017) gives a more reliable estimate of the right amount of clusters. In addition, Berghuisj et al (2014) also found that 10 clusters are a good number to capture the differences in hydrology in the continental US. However, this number of clusters still remains discussable, which is why we provide in the revised version of the manuscript the choice of our number of clusters (section 3.4).

- Is the distribution between clusters so skewed because the catchment sample is not uniformly distributed across the selected attributes?

We think the skewed distribution can be explained quite well with the distance in PCA space: The small clusters are just very different to the bigger clusters (Figure 3). We see no reason, why hydrological behavior in rivers should be uniformly distributed. Many other kinds of natural phenomenon are shaped in bell curves. For example the size of humans is based on a large set of factors and follows a bell curve. The extreme ends of the curve can only be reached if many factors align in the same direction. This is probably true for rivers as well. "Normal" behavior can probably be reached by all kinds of combinations of catchment attributes, while more extreme behavior needs to have several attributes that force it into that direction.

- Would more and/or different attributes provide more balanced clustering results?

The clustering is not based on the catchment attributes, only on the hydrological signatures.

- If the catchment sample is not uniformly distributed across attribute space, does this influence the PCA results?

The attributes do not affect the clustering and with this also not the PCA results.

L126. That aridity and forest fraction score highest could possibly relate to the high correlation between these two attributes. Investigating the correlations between the 15 catchment attributes could show how much independent information is contained in each. The same could be said about fractional snowfall and elevation. – Note: upon further reading I see that these correlations are in Figure 6. This information should be part of the text here.

We now discuss the correlations earlier (see answer to main point 1).

L137. I don't think calling these six signatures "more hydrologically meaningful" is supported by the findings of Addor et al. (2018). "more gradually varying in space" perhaps.

Changed as proposed.

L144. "This can probably be extrapolated to most catchments in the continental US without human influence, as the CAMELS dataset contains large samples of undisturbed catchments". This sentence is speculation and should be removed. If the authors want to keep this statement it could for example be supported by calculating the climate attributes used by Knoben et al. (2018) and comparing these to the range of values for these attributes found across the CONUS. This would show how climatically representative the CAMELS catchments are for the wider CONUS.

We removed this sentence.

L185. See also Berghuijs et al. (2014) who find hydrologic similarity across comparable distance in the CONUS; or Kuentz et al. (2017) who find hydrologic similarity across comparable distances in Europe;

or Knoben et al. (2018) who find catchments with similar hydrologic regime on different continents, using only climate indicators to describe similarity.

We reference those studies now and mention their similar findings.

L189-195. I suspect that if correlations between attributes are taken into account, many of the attributes that are of high importance in each cluster turn out to be quite directly related one another. For example, (cluster 1) high aridity and low forest fraction & green vegetation fraction maximum will be inversely correlated; (cluster 3) precipitation and snow and elevation will be correlated, as will forest fraction and LAI maximum and green vegetation fraction maximum. Therefore I expect that this part of the analysis will be more instructive if these correlations are accounted for, either in selection of the attributes or by lumping correlated attributes into groups in some fashion. Changes to Figure 5 might be needed.

We discuss the correlations now at several places throughout the manuscript.

L214. "While aridity … single clusters (Figure 5)." Implying that aridity is not important in most of the clusters seems a bit of a stretch. Aridity is the most important attribute in 4 out of 10 clusters, and the second-most important in another 2. It appears in the top 5 of important attributes in 8 out of 10 clusters (and in the remaining 2 clusters the correlated forest fraction appears), more often than any other attribute.

We have rewritten this whole section (3.5), see earlier answers.

L248. "Therefore, our selection of hydrological signatures seems to allow a better identification of hydrological similarities." Unfortunately I think this argument can be reversed as well, in the sense that this selection of signatures might not capture enough of the details of the individual regimes to give the clustering approach any trouble. Because these 6 signatures are strongly related to climate (e.g. Addor et al., 2018; Knoben et al. 2018), and the relevant climate indices are (mostly) included in the clustering approach, it is not surprising that these signatures cluster easily. The fact that the authors don't use a climate seasonality attribute, which has been shown to be an important driver of hydrologic differences, could potentially explain why their Cluster 2 does not seem to have any distinct character. Instead of making this statement and moving on, a strong contribution would be if the authors can determine how to make hydrologic sense of all the catchments that don't seem to follow any obvious pattern. Would different attributes solve this?

We included the climate seasonality metric given in the CAMELS data set and discuss it in sections 3.4 and 3.5. This improved the distinctions for some of the cluster. However, for the mentioned cluster 2 the climate seasonality did not provide much additional information.

L301. I'd argue that Cluster 4 seems to be firmly placed in the non-arid & snow-dominated region of the climate space. There are more catchments in this climate region that belong to different clusters but this is (1) inherent to imposing binary boundaries (catchments are either cluster X or Y, even if they are 49% similar to X and 51% similar to Y) and (2) because the climate plots in Figure 8 only look at a limited selection of possibly influential attributes (climatic or otherwise).

The reviewer is right. Accordingly, we removed the sentence.

L310-315. This connection between signatures and climate can also be seen in Knoben et al. (2018) and Kuentz et al. (2017). Addor et al. (2018), Knoben at al. (2018) and Kuentz et al. (2017) (among others) acknowledge that using climate alone is not sufficient to produce a catchment classification system. This should probably be mentioned as part of this section (or in the introduction of the paper, because it provides a compelling reason for investigating catchment attributes).

We reference this now at the end of the mentioned chapter.

Figure 8. Is the aridity axis upside down in these plots? More arid catchments seem to have higher flows.

It seems that we accidently switched the sign of the aridity values during the extraction from the climate maps of Knoben et al. (2018). We fixed this and the figure should be correct now.

Figure A2. I like the way violin plots look, but kernel density smoothing does not respect physical boundaries very well and distorts the data being plotted. See for example cluster 3 and the mean winter discharge signature, which is, according to the violin, a negative flux for some of the catchments in this cluster. Histograms or box-and-whisker plots would more accurately reflect the data.

As recommended by the reviewer, we now display this information as boxplots.

Figure A3. See comment above.

Changed as proposed.

Figure A4. Is the aridity axis upside down in these plots? More arid catchments seem to have higher flows.

Fixed.

Typographical

30. “those” > “this”?

119. Kuentz et al. (2018) > Kuentz et al. (2017)

P11. Caption of Figure 5. “For the catchment clusters.” should not be a stand-alone sentence.

215. “single” > “individual”?

251. I understand what this sentence is meant to say but it doesn’t quite work. Is “This human influence might mask otherwise apparent patterns.” better?

261. “have” > “has”

265. “cluster” > “clusters”

463. "… the low elevation those catchments are located, …" > "… the low elevation those catchments are located at, …"

500. "cluster catchments". Should the word "catchments" be here?

All changed as proposed.

In addition to the changes mentioned above, we also rewrote parts of the abstract and the summary to accommodate for the changes made on the manuscript.

Abstract:

*The behavior of every catchment is unique. Still, we seek for ways to classify them as this helps to improve hydrological theories. In this study, we use hydrological signatures that were recently identified as those with highest spatial predictability to clusters 643 catchments from the CAMELS data set. We describe the resulting clusters concerning their behavior, location and attributes. We then analyze the connections between the resulting clusters and the catchment attributes and relate this to the co-variability of the catchment attributes in the eastern and western US. To explore whether the observed differences result from clustering catchments by either climate or hydrological behavior, we compare the hydrological clusters to climatic ones. We find that for the overall data set climate is the most important factor for the hydrological behavior. However, depending on the location, either aridity, snow or seasonality has the largest influence. The clusters derived from the hydrological signatures partly follow eco regions in the US and can be grouped into four main behavior trends. In addition, the clusters show consistent low flow behavior, even though the hydrological signatures used describe high and mean flows only. We can also show that most of the catchments in the CAMELS dataset have a low range of captured hydrological behaviors, while some, more extreme catchments, derivate form that trend. In the comparison of climatic and hydrological clusters, we see that the widely used Koeppen-Geiger climate classification is not suitable to find hydrologically similar catchments. However, in comparison with a novel, hydrologically based continuous climate classifications, some clusters follow the climate classification very directly, whilst others do not. From those results, we conclude that the signal of the climatic forcing can be found more explicitly in the behavior of some catchments than in others. It remains unclear if this is caused by a higher intra-catchment variability of the climate or a higher influence of other catchment attributes, overlaying the climate signal. Our findings suggest that very different sets of catchment attributes and climate can cause very similar hydrological behavior of catchments - a sort of equifinality of the catchment response.*

Summary and conclusions

*This study explored differences in the catchment characteristics between the eastern and western US, the properties and location of catchment clusters based on hydrological signatures, the importance of catchment attributes for those clusters and how this study relates to other clustering studies and methods. We found that the correlations catchment characteristics are quite similar for the eastern and western US with the exception of mean elevation, snow, geology and the leaf area index. For the overall CAMELS data set climate seems to be the most important factor for the hydrological behavior. However, depending on the location either aridity, snow or seasonality were most important. The clusters derived from the hydrological signatures partly follow the eco regions in the US and can be combined into four groups of general behavior trends. Still, similar catchments can be quite far away from each other. We also found that most of the catchments have a rather similar discharge* *behavior, while only some, more extreme catchments, derivate from that main trend. This might be a hint why it is so difficult to clusters catchments, as those single extreme catchments are quite unique and do not fit together well with other catchments. We also found, that there are differences of how directly the signal of forcing climate can be found again in the hydrological behavior. This explains why catchments often show a surprisingly similar behavior across many different climate and landscape properties (Troch et al., 2013) and why the most hydrologically similar catchment can be hundreds of kilometers away. Those findings also relate to the paradox that small scales/single catchment studies identify geology/soils as most important for the hydrological behavior, while large sample studies usually find the climate as most important. This might simply be influenced by spatial proximity. Small scale studies look at catchment which all have a similar climatic forcing and thus only the other catchment attributes can be the cause of differences in hydrological behavior. Large sample studies on the other hand consider catchment from a wider area and thus attribute the differences in behavior to climate.*

*The aggregated data used in this study might level out the variability of the catchment attributes in the single catchment, but it also indicates that there is a kind of equifinality in the behavior of catchments. Different sets of intertwined climate forcing and catchment attributes could lead to a very similar overall behavior, not unlike to hydrological models that produce the same discharge with different sets of parameters.*

*We acknowledge that the results are dependent on the amount and size of the clusters, the catchment attributes considered and the hydrological signatures used. Still, we think that the CAMELS dataset offers an excellent overview of different kinds of catchments in contrasting climatic and topographic regions. In addition, this study shows that using hydrological signatures with high spatial predictability results in hydrological meaningful clusters, which show consistent low flow behavior, even though those low flows were not explicitly considered. However, it seems that even a comprehensive dataset like CAMELS, does not allow an easy way to find a conclusive set of clusters for catchments. For future research, we recommend to include measures of spatial variability of the climate in the single catchments and to look into the single clusters in more depth. This might help to prove, if a less clear climatic signal is caused by intra-catchment variability of the climate or a larger influence of other catchment attributes.*

# Using hydrological and climatic catchment clusters to explore drivers of catchment behavior

Florian U. Jehn[1], Konrad Bestian[1], Lutz Breuer[1,2], Philipp Kraft[1], Tobias Houska[1]

[1]Institute for Landscape Ecology and Resources Management (ILR), Research Centre for BioSystems, Land Use and Nutrition (iFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26, 35390 Giessen, Germany
[2]Centre for International Development and Environmental Research (ZEU), Justus Liebig University Giessen, Senckenbergstrasse 3, 35392 Giessen, Germany

*Correspondence to*: Florian U. Jehn (florian.u.jehn@umwelt.uni-giessen.de)

**Abstract.**

The behavior of every catchment is unique. Still, we seek for ways to classify them as this helps to improve hydrological theories. In this study, we use hydrological signatures that were recently identified as those with highest spatial predictability to clusters 643 catchments from the CAMELS data set. We describe the resulting clusters concerning their behavior, location and attributes. -We then~~We~~ analyze the connections between the resulting clusters and the catchment attributes and relate this to the co-variability of the catchment attributes in the eastern and western US. To explore whether the observed differences result from clustering catchments by either climate or hydrological behavior, we compare the hydrological clusters to climatic ones. We find that for the overall data set climate is the most important factor for the hydrological behavior. However, depending on the location, either aridity, snow or seasonality has the largest influence. The clusters derived from the hydrological signatures partly follow eco regions in the US and can be grouped into four main behavior trends. In addition, the clusters show consistent low flow behavior, even though the hydrological signatures used describe high and mean flows only. We can also show that most of the catchments in the CAMELS dataset have a low range of hydrological behaviors, while some, more extreme catchments, derivate form that trend. ~~We find that aridity is more important for hydrological behavior in the eastern US, while it is the amount of snow in the West.~~ In the comparison of climatic and hydrological clusters, we see that the widely used Koeppen-Geiger climate classification is ~~unsuitable~~ not suitable to find hydrologically similar catchments. However, in comparison with a novel, hydrologically based continuous climate classifications, some clusters follow the climate classification very directly, whilst others do not. From those results, we conclude that the signal of the climatic forcing can be found more explicitly in the behavior of some catchments than in others. It remains unclear if this is caused by a higher intra-catchment variability of the climate or a higher influence of other catchment attributes, overlaying the climate signal. Our findings suggest that very different sets of catchment attributes and climate can cause very similar hydrological behavior of catchments - a sort of equifinality of the catchment response.

## 1 Introduction

Every hydrological catchment is composed of a unique combination of topography and climate, which makes their discharge heterogeneous. This, in turn, makes it hard to generalize behavior beyond individual catchments (Beven, 2000). Catchment classification is used to find patterns and laws in the heterogeneity of landscapes and climatic inputs (Sivapalan, 2003). Historically, this classification was often done by simply using geographic, administrative or physiographic considerations. However, those regions proved to be not sufficiently homogenous (Burn, 1997). Therefore, it was proposed to use seasonality measures with physiographic and meteorological characteristics, but it was deemed difficult to obtain ~~those~~ this information for a large number of catchments (Burn, 1997), even if only simple catchment attributes (e.g. aridity) are used (Wagener et al., 2007). Nonetheless, in the last decade datasets with hydrologic and geological data were made available, comprising information of hundreds of catchments around the world (Addor et al., 2017; Alvarez-Garreton et al., 2018; Newman et al., 2014; Schaake et al., 2006). This is a significant step forward as those large sample datasets can generate new insights, which are impossible to obtain when only a few catchments are considered (Gupta et al., 2014). Different attributes have been used to classify groups of catchments in those kind of datasets: flow duration curve (Coopersmith et al., 2012; Yaeger et al., 2012), catchment structure (McGlynn and Seibert, 2003), hydro-climatic regions (Potter et al., 2005), function response (Sivapalan, 2005) and more recently, a variety of hydrological signatures (Kuentz et al., 2017; Sawicz et al., 2011; Toth, 2013). Quite often, climate has been identified as the most important driving factor for different hydrological behavior (Berghuijs et al., 2014; Kuentz et al., 2017; Sawicz et al., 2011). Still, it is also noted that this does not hold true for all regions and scales (Ali et al., 2012; Singh et al., 2014; Trancoso et al., 2017). In addition, a recent large study of Addor et al. (2018) has shown that many of the hydrological signatures often used for classification, are easily affected by data uncertainties and cannot be predicted using catchment attributes. Another recent study by Kuentz et al. (2017) used an extremely large datasets of 35,000 catchments in Europe and classified them using hydrological signatures. For their classification, they used hierarchical clustering and evaluated the result of the clustering by comparing variance between different numbers of clusters. They were able to find ten distinct classes of catchments. However, Kuentz et al. (2017) used some of the signatures identified to have a low spatial predictability by Addor et al. (2018). In addition, one third of their catchments was aggregated in one large class with no distinguishable attributes. Overall, we conclude that no large sample study exists that uses only hydrological signatures with a good spatial predictability. In addition, if the climate is the dominant driver of catchment behavior, clustering catchments based on their hydrological behavior should result in clusters with a uniform climate.

Therefore, we selected the best six hydrological signatures with spatial predictability to classify catchments of the CAMELS (Catchment Attributes and MEteorology for Large-Sample Studies) dataset (Addor et al., 2017). Those six hydrological signatures are evaluated together with the ~~fif~~sixteen catchment attributes that were shown to have a large influence on hydrological signatures (Addor et al., 2018). The connection between the hydrological signatures and the catchment attributes is determined by using quadratic regression of the principal components (of the hydrological signatures) and the catchment attributes. This will help to explore, if a clustering with hydrological signatures that have a high predictability in space, provides hydrologically meaningful clusters and how those are related to catchment attributes. In addition, we compare the hydrologically derived clusters with climatic clusters and determine the spatial distance between the most hydrologically similar catchments. This will determine if grouping catchments by climate or by hydrologic behavior will yield the same results and if the signatures identified by Addor et al. (2018) as having the highest spatial predictability can be used to delineate hydrologically meaningful clusters, even though they do not consider low flows. ~~explore the validity of considering spatial distance as a measure of similarity between catchments.~~

## 2 Material and Methods

### 2.1 Data base

This work is based on a detailed analysis of catchment attributes and information contained in hydrological signatures. The CAMELS data set contains 671 catchment in the continental united states (Addor et al., 2017) with additional meta information such as slope and vegetation parameters. For our study, we used a selection of the available meta data. We excluded all catchments that had missing data, which left us with 643 catchments. Those catchments come from a wide spectrum of characteristics like different climatic regions, elevations ranging from 10 to almost 3,600 m a.s.l. and catchment areas ranging from 4 to almost 26,000 km². ~~To ensure an equal representation of the different catchment attributes classes (climate, topography, vegetation, soil, geology) we~~We used ~~three~~ the following attributes per class. *Climate*: aridity, frequency of high precipitation events, fraction of precipitation falling as snow; precipitation seasonality, *Vegetation*: forest fraction, green vegetation fraction maximum, LAI maximum; *Topography*: mean slope, mean elevation, catchment area; *Soil*: clay fraction, depth to bedrock, sand fraction; *Geology*: dominant geological class, subsurface porosity, subsurface permeability. Those catchment attributes were chosen due to their ability to improve the prediction of hydrological signatures (Addor et al., 2018) and because they are relatively easy to obtain, which will allow a transfer of this method to other groups of catchments world-wide.

Hydrological signatures cover different behaviors of catchments. However, many of the published signatures have large uncertainties (Westerberg and McMillan, 2015) and lack in predictive power (Addor et al., 2018). Therefore, we used the six hydrological signatures with the best predictability in space (Table 1) (Addor et al., 2018). Those signatures were calculated for all catchments. Due to this selection, no signatures that capture low flow behavior were used, as those signatures have a very low spatial predictability.

**Table 1: Applied hydrological signatures on the discharge data of the CAMELS data set (Addor et al., 2018).**

| Signature | Unit |
|---|---|
| Mean annual daily discharge | mm $d^{-1}$ |
| Mean winter daily discharge (Nov. – Apr.) | mm $d^{-1}$ |
| Mean half-flow date; Date on which the cumulative discharge since October first reaches half of the annual discharge | day of year |
| 95 % Flow quantile (high flow) | mm $d^{-1}$ |
| Runoff ratio | - |
| Mean summer daily discharge (May – Oct.) | mm $d^{-1}$ |

### 2.2 Data analysis

The workflow of the data analysis considers a data reduction approach with a principal component analysis and a subsequent clustering of the principal components, similar to Kuentz et al. (2017) and McManamay et al. (2014). For the principal component analysis and the clustering, we used the Python package sklearn (0.19.1). The code is available at GitHub (Jehn, 2018). Validity was checked by a random selection of 50 and 75 % of all catchments. We found that the overall picture stayed the same (not shown). In all further analysis, we used all catchments to get a sample as large as possible to be able to make statements that are more general.

*Calculation of the principal component analysis*

The principal components were calculated from the six hydrological signatures described above (Table 1). We used a principal component analysis on the hydrological signatures to remove correlations between the single hydrological signatures. We only used principal components that together account for at least 80% of the total variance of the hydrological signatures, which resulted in two principal components. Those two principal components contain the uncorrelated information of all hydrological signatures used and thus can be seen as describers of the ~~overall~~ hydrological behavior in regard to the overall amount of discharge, its distribution throughout the year, high flows and runoff-ratio. Therefore, catchments with similar principal components have similar hydrological behavior along those signatures.

*Evaluating the connection between the principal components and the catchment attributes*

1) First, we calculated quadratic regressions between the two principal components and the catchment attributes (with the principal component as the dependent variable). This resulted in one coefficient of determination ($R^2$) for each pair of principal component and catchment attribute (e.g. PC 1 and aridity).

2) We then weighted the $R^2$ by the explained variance of the principal components. This addresses the differences in the explained variance of the principal components (e.g., PC 1 explained 75% of the variance, PC 2 explained 19% of the variance).
3) The weighted coefficients of determination of the two principal components were subsequently added to obtain one coefficient of determination for every catchment attribute.

Quadratic regression was selected as interactions in natural hydrological systems are known to have unclear patterns and can therefore often not be fitted with a simple straight line (Addor et al., 2017; Costanza et al., 1993). This was done first for the whole dataset and then for all clusters separately. This procedure captures the pattern on the catchment attributes in the PCA space of the hydrological signatures (for examples of this pattern see Figure A1).

*Clustering the principal components*

The principal components of the hydrological signatures were clustered following agglomerative hierarchical clustering with ward linkage (Ward, 1963), similar to previous studies (Kuentz et al., 2017; Li et al., 2018; Yeung and Ruzzo, 2001). Therefore, the clusters are based on the hydrological signatures of the catchments. From ~~those~~ the previous studies, Kuentz et al. (2017~~8~~) provides the largest set with over 35,000 catchments. They also clustered their catchments in a PCA space of a range of hydrological signatures. ~~-~~To select the number of clusters, they used the elbow method (and two other methods to validate their results) and found that ten or eleven clusters (depending on the method) were most appropriate for their data. Due to the similarity in the clustered data and the larger database of Kuentz et al. (201~~8~~7), we also used ten clusters. (Berghuijs et al., 2014) also found that ten clusters captured the distinct hydrological behaviors for the continental US. Those ten clusters represent groups of catchments with distinctly different hydrological behavior.

## 3 Results and Discussion

### 3.1 Catchment attribute correlations in the CAMELS data set

Usually the $100^{th}$ ~~100th~~ meridian is seen as the dividing climatic line in the US, splitting the country in a semi-arid west and a humid east. We assume that this difference in climate also has implications for the hydrology and the overall catchment attributes in those regions. To quantify this we split the CAMELS data set into a western and an eastern part, based on the 100th meridian (Figure 1 and 4). This shows that many of the catchment attribute correlations do not differ much between the east and the west. In most cases (>80%), Pearson correlation coefficients vary by less than 0.4 (Figure 1c). Still, there are some catchment attributes with larger differences of up to 0.7 between both regions. Most striking are the mean elevation and the fraction of the precipitation falling as snow as well as the vegetation attributes LAI maximum and Green vegetation fraction maximum. Even though these attributes are directly related to each other through temperature gradients, they differ substantially in both parts of the country. In the mountainous western US, elevation is highly correlated with the fraction of precipitation falling as snow (r=0.8), while it is not in the eastern US (r=0.4). This, and the different correlations between vegetation and elevation are probably caused by the fact that the temperature gradients differ in both regions. In the western US it is much more mountainous and thus temperatures typically change with elevation. In the more level eastern US, on the other hand, the change in temperature is mainly linked to the latitude. Striking are also the changes of correlation with regard to the fraction of precipitation falling as snow. Here we find altered directions of the correlation, i.e., positive correlations with LAI maximum and frequency of high precipitation events in the east turn to negative ones in the west. It also becomes obvious that all three measures of vegetation seem to track similar characteristics in the catchments, as they highly correlate with each other (especially in the eastern US with r=0.9). In addition, all vegetation attributes depict a large negative correlation with aridity. Hence, the vegetation attributes considered are likely good proxies for aridity. Overall, we see that the relations between the catchment attributes are quite similar for the eastern and western US, with the exception of the mean elevation, snow and the LAI maximum.

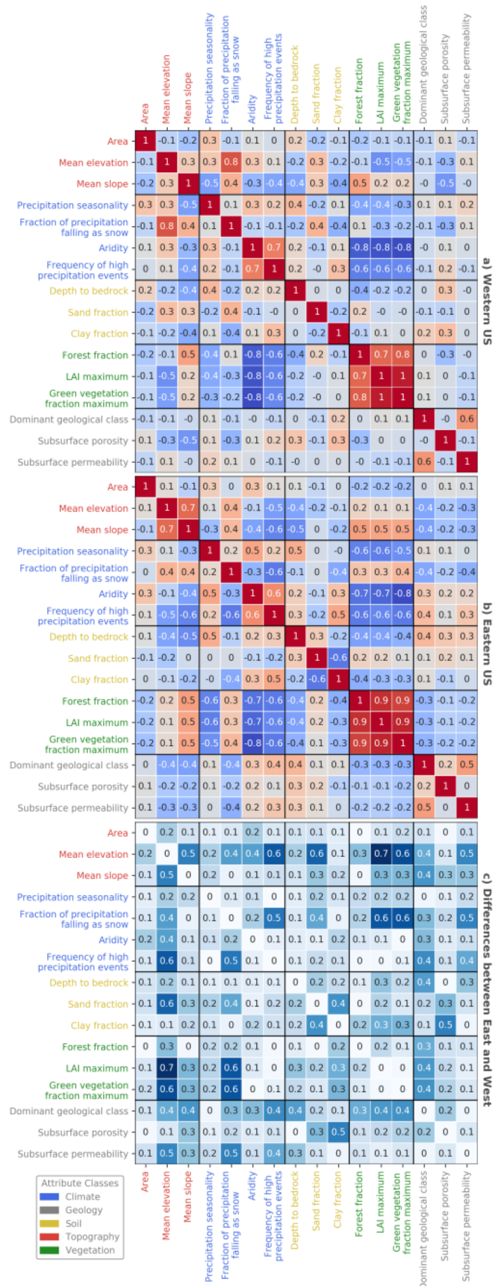



**Figure 1~~6~~: Pearson ~~C~~correlation coefficients given for~~of~~ all catchment attributes in~~for~~ western (a) and eastern (b) US. ~~and a~~Absolute differences~~ c) in~~ of the correlation coefficients between the eastern and western US is given in c). Eastern and western is defined by the 100th meridian. Due to rounding effects, correlations with the same Pearson correlation coefficient might show slightly varying color codes.**

### 3.~~1~~2 Impacts of catchment attributes on discharge characteristics in the whole dataset

~~First we~~Next we examined the weighted $R^2$ of the catchment attributes for the whole dataset. This analysis shows not only differences in their score between the single attributes, but also between the different classes of catchment attributes (Figure ~~1~~2). Attributes related to climate (aridity) and vegetation (forest fraction) get the highest scores. With the exception of the mean slope, the first seven catchment attributes are all related to climate and vegetation. The last seven attributes on the other hand are all related to soil and geology, except the catchment area. They also show much lower scores of the weighted $R^2$. This indicates that soil and geology are less important for the chosen hydrological signatures. Similar patterns were also found by (Yaeger et al., 2012). They stated climate as the most important driver for the hydrology. As the correlations between the catchment attributes showed that the climate and the vegetation attributes are highly correlated (Figure 1), it can be assumed that climate is the overall most important factor, with aridity and high precipitation events being most important.

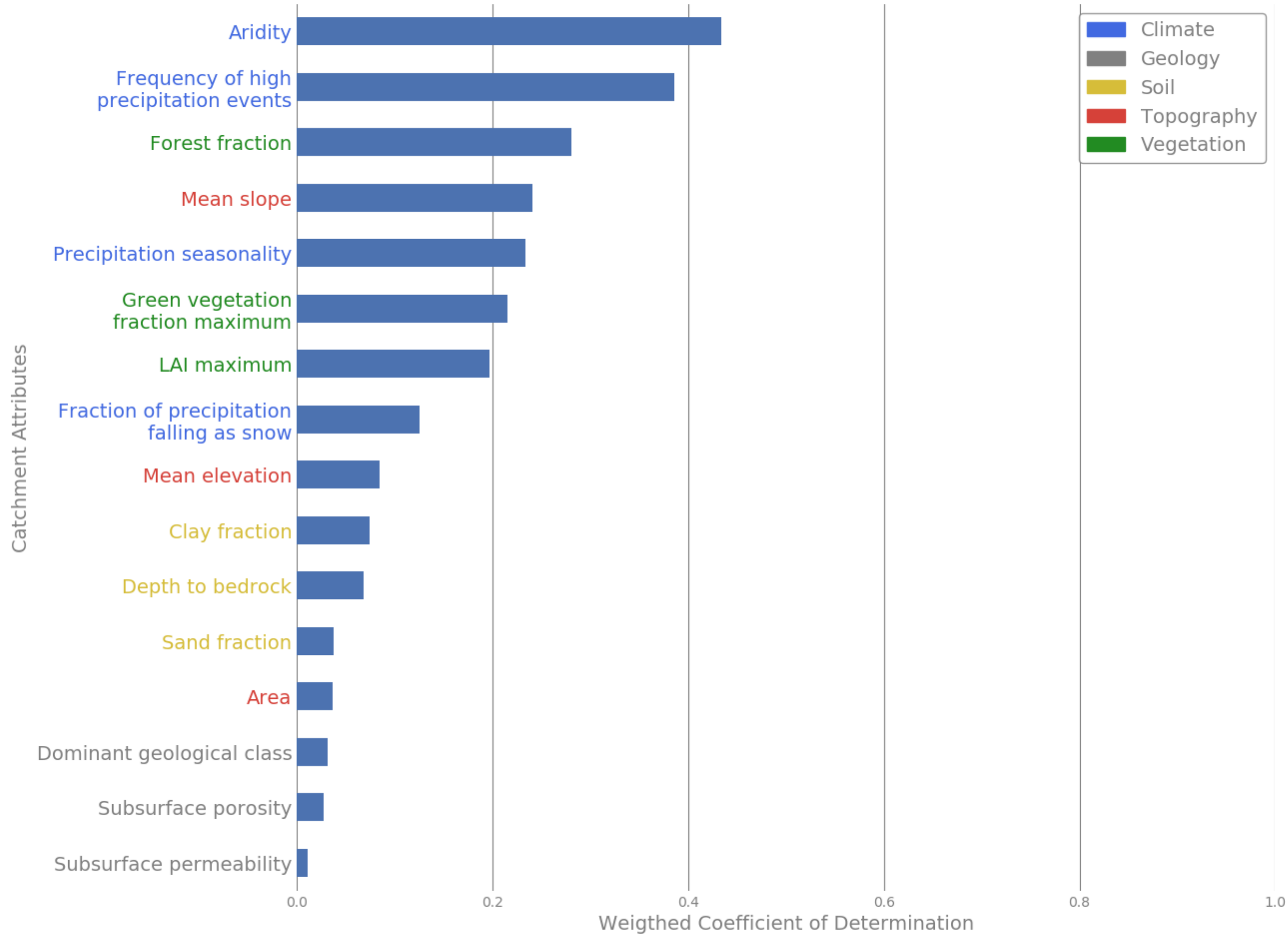



**Figure ~~1~~2: Importance of catchment attributes evaluated by quadratic regression for all considered catchments. Attributes colored according to their catchment attribute class.**

However, they also unraveled that low flows are mainly controlled by soil and geology. The minor importance of soil and geology in our study might therefore be biased by the choice of hydrological signatures, which excluded low flow signatures due to their low predictability in space. Nevertheless, our study probably captures a more general trend as we used a larger dataset and ~~more hydrologically meaningful~~ hydrological signatures that vary more gradually in space (Addor et al., 2018). Addor et al. (2018) also explored the influence of different catchment attributes in the CAMELS dataset on discharge characteristics. They found that climate has the largest influence on discharge characteristics, well in agreement with Coopersmith et al. (2012). The latter also used a large group of catchments in the continental United States from the MOPEX dataset. They conclude that the seasonality of the climate is the most important driver of discharge characteristics, which seems to be less important for the overall data set in our analysis. However, Coopersmith et al. (2012) only analyzed the flow duration curve, which has a mediocre predictability in space and it is therefore ~~more unclear~~less clear what it really depicts (Addor et al., 2018). Overall, this study here is in line with other literature in the field. Using the weighted R² reliably detects climatic forcing as the most important for the discharge characteristics for a large group of catchments. ~~This can probably be extrapolated to most catchments in the continental US without human influence, as the CAMELS dataset contains large samples of undisturbed catchments (Addor et al., 2017). In the next step, we will test whether these relations also hold for the clusters of the catchments.~~

### 3.~~2~~3 Relation of the principal components and the hydrological signatures

The rivers considered in this study show a wide range in hydrological signatures. This is visible in the clusters of principal components of the hydrological signatures (Figure ~~2~~3).Most of the rivers are opposite of the loading vectors (the loading vectors are shown as arrows). This shows that most rivers have relatively low values for all hydrological signatures and only some, more extreme rivers, have higher values for specific hydrological signatures. Most typical for the overall behavior of the river are the hydrological signatures mean annual discharge and Q95 (high flows), as they have a strong correlation with the first principal component. For the second principal component, the mean half-flow date ~~(an indicator for seasonality)~~ has the highest correlation. Therefore, the first principal component can be seen as a measure of overall discharge and amount of high flows.~~, while the second principal component can be seen as a measure of seasonality in the discharge.~~ Overall, it can also be seen that most of the rivers show a relatively similar behavior (cluster 1, 2, 8, 9, 10), while smaller groups of rivers tend to derivate from that by having a more extreme behavior (cluster 3, 5, 7). The remaining clusters 4 and 6 are located between those extremes. This pattern also explains the different sizes of the clusters. While most catchments behave relatively similar, only some show extreme behavior and thus the clusters with extreme catchments are smaller.

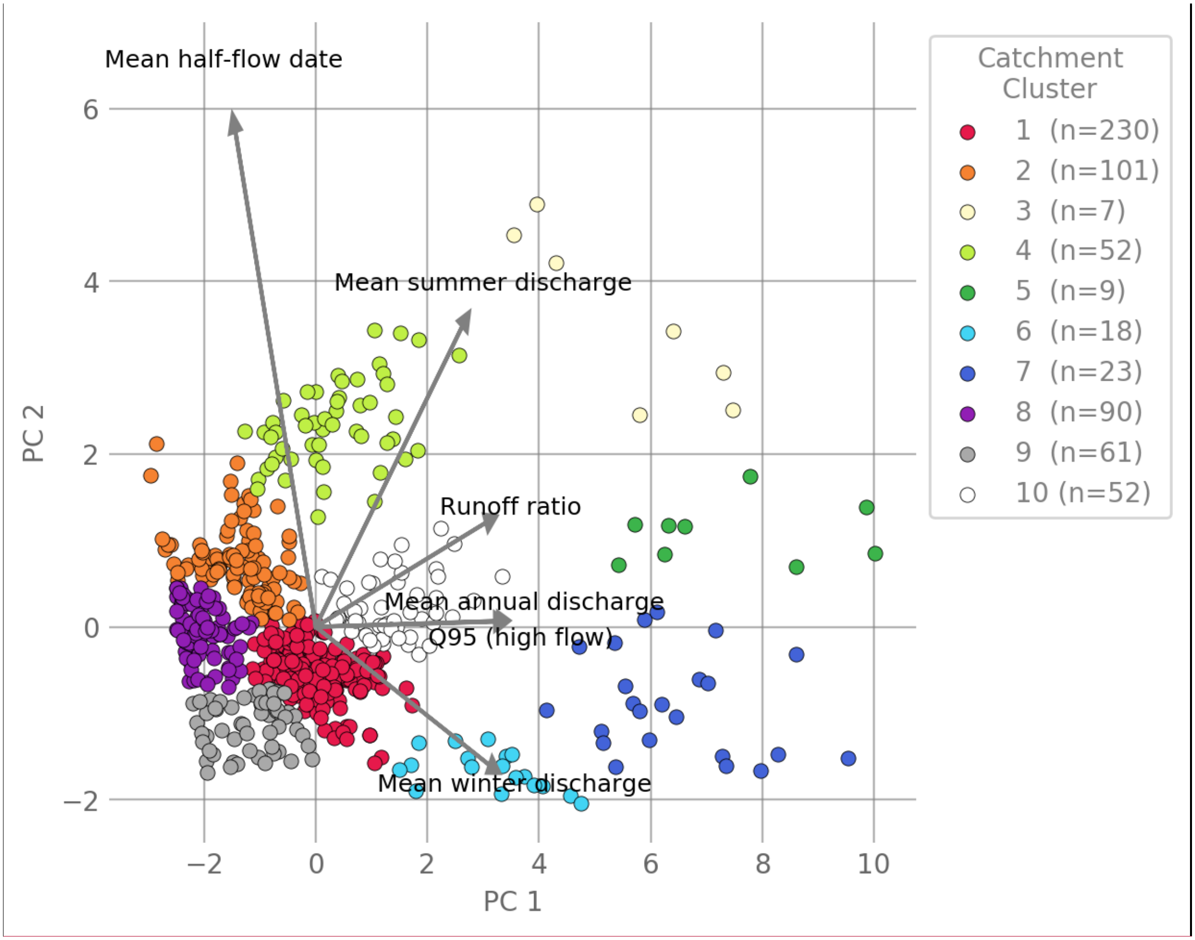



**Figure ~~2~~3: Biplot of the principal components (PC). Colors indicate the cluster of the catchment.**

### 3.4~~3~~ Location and properties of the catchment clusters ~~Exploration of the catchment clusters~~

The catchment attributes in the CAMELS and similar large scale datasets often show a pattern that resembles climatic zones (Addor et al., 2018; Coopersmith et al., 2012; Yaeger et al., 2012). ~~The picture is less clear for the hydrological catchment clusters presented. This is directly observable in the spatial distribution of the clusters (Figure 3).~~ For the catchments clusters presented here, we can see that most of the clusters roughly follow ecoregions in the US (Figure 4). Especially clusters 1, 4, 6 and 7 are almost entirely located within one ecoregion. Cluster 2, 8 and 9 on the other hand follow those ecological boundaries to a lesser degree.

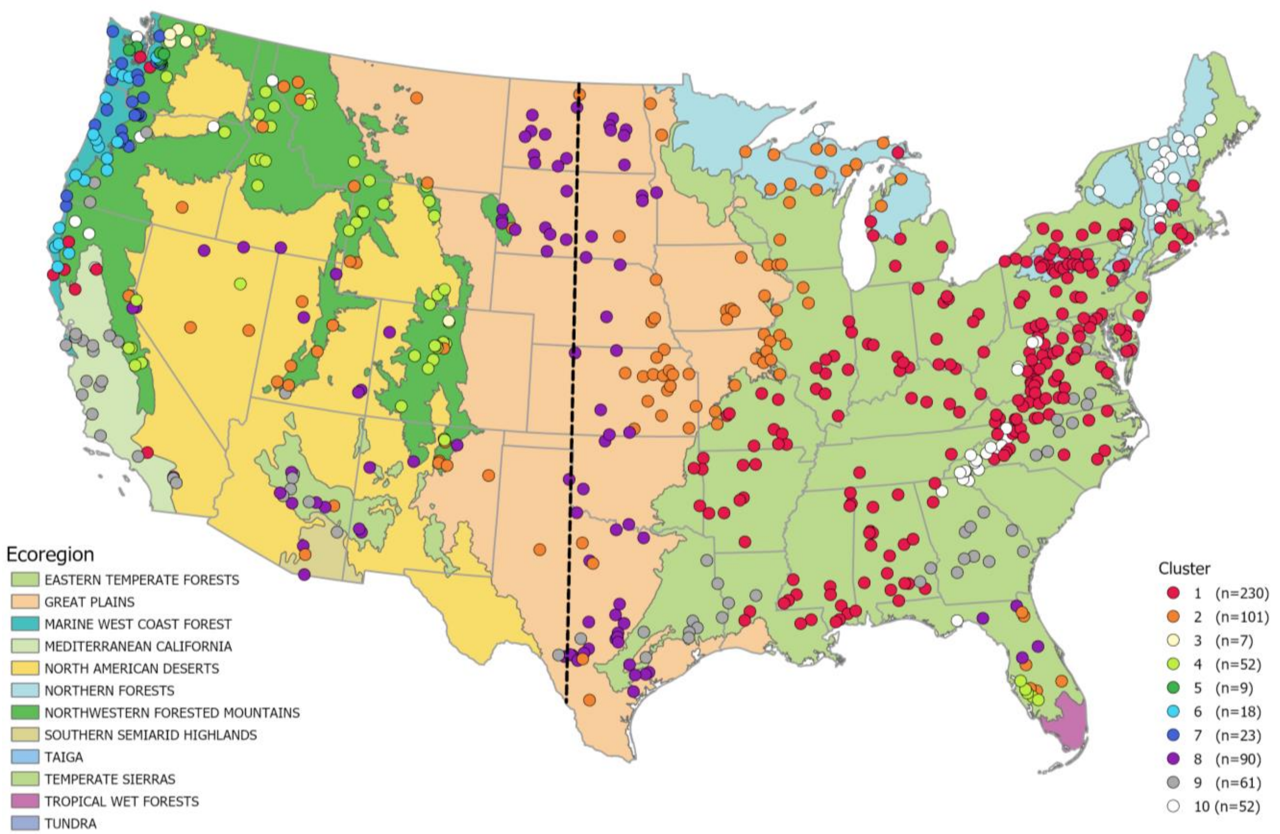



**Figure ~~3~~4: Locations of the clustered CAMELS catchments and level I ecoregions (Omernik and Griffith, 2014) in the continental US. Dotted line marks the 100th meridian.**

We can see a split of the clusters along the 100th meridian. ~~This split can also be found in some of the clusters depicted in Figure 3.~~ Cluster 3, 4, 5, 6 and 7 are ~~all~~ located mainly in the ~~W~~west, while Cluster 1 and 10 are mainly found in the ~~E~~east. However, the remaining Clusters 2, 8 and 9 have roughly similar ~~amounts~~ numbers of catchments in both regions. Overall, ~~T~~the catchments in the eastern half of the United States form large spatial patterns of similar behavior, while the catchments in the west are ~~a lot~~ patchier. ~~The descriptions of the catchment clusters are summarized in Table 2. A further detailed description of the clusters can be found in the appendix, together with figures showing the distribution of hydrological signatures (Figure A2) and catchment attributes (Figure A3) in the clusters.~~ ~~A list of all catchments with index, position and cluster classification is given in the supplementary material.~~


**Figure 5: Swarm plot of the real world distances of all catchments to the most hydrologically similar catchment (based on their distance in the PCA space of the hydrological signatures).**

In addition, similar catchments can be quite far away from each other (Figure 5). Sometimes, the catchment with the most similar signature was found as far as 4,000 km away (almost the entire longitudinal distance of the continental US). This explains why spatial proximity seems to be important in some studies that look into explanations of catchment behavior (Andréassian et al., 2012; Sawicz et al., 2011), but not in others (Trancoso et al., 2017). This also indicates that clustering by using spatial proximity might only work in regions like the eastern US, where the behavior of rivers changes only gradually. The finding that the most similar catchment (based on their hydrological signatures) can be far away, also explains the behavior of clusters that contain catchments quite distant from each other (e.g. Cluster 4). Even though the catchments might be far away from each other, the interplay of different catchment attributes and driving factors, including sometimes very different climates, can lead to similar (equifinal) discharge behavior, concerning the overall amount of discharge, its distribution in the year, the high flows and the runoff-ratio. This was also found by several other studies (e.g. (Berghuijs et al., 2014; Knoben et al., 2018; Kuentz et al., 2017)).

In the following, we describe the catchment clusters in regard to their characteristics in meteorology (Figure 6), attributes (Figure 7), hydrology (Figure 8) and location (Figure 4). The main points of this description are summarized in Table 2. A list of all catchments with index, position, ~~and~~ cluster classification and climate indices is given in the supplementary material.

**Cluster 1** is defined by a dense vegetation cover (Figure 7). The low elevation of those catchments results in little annual snow fall. They are mainly located in the southeastern and central plains and therefore get relative high rainfall (>1,000 mm per year) (Figure 4), almost uniformly distributed over the year (Figure 6). From a hydrological perspective, these catchments produce little discharge. Cluster 1 contains the highest number of catchments (n=230). So over one third of the catchments in CAMELS show a relatively similar behavior when it comes to the amount of water fluxes and their distribution throughout the year. This is particular visible when we look at annual supply of discharge (Figure 6). Even though the cluster contains a large number of catchments that also partly differ a lot in their potential evapotranspiration, there is only a minor difference in the amount of discharge and its seasonality.~~is defined by a high cover of vegetation. In addition, most catchments are located~~

~~at low elevations, experience little snow and have a deep bedrock. Hydrologically these catchments have little discharge. They are mainly located in the Southeastern and Central Plains and therefore get relative high rainfall (> 1000 mm year). Their low discharge is probably caused by the low elevation those catchments are located, groundwater discharge and the high evaporation of the forests. Cluster 1 also contains the largest amount of catchments from all cluster (n = 230). So over one third of the catchments in CAMELS show a relatively similar behavior.~~

**Cluster 2**'s most typical attribute is its high precipitation seasonality. However, concerning most other catchment attributes, Cluster 2 is undefined as it contains catchments of most regions of the continental US (with a concentration in the eastern Great Plains) (Figure 4). The hydrological signatures on the other hand show a clearer pattern. Here, the mean winter discharge, Q95 and the mean annual discharge have a narrow range (Figure 8). This shows that catchments with very different attributes can produce similar discharge characteristics. The different attributes seem to cancel each other out in their influence on the discharge. This might be enhanced by the high precipitation seasonality with higher precipitation in the summer, which creates a strong climatic forcing and thus a narrow range for the hydrological signatures (Figure 6). This cluster differs from the first one, by having even lower discharge, with almost no peaks and a higher influence of snow melt. ~~Cluster 2 most typical attribute in comparison with the other catchments is its depth to the bedrock. However, concerning the catchment attributes cluster 2 is undefined as it contains catchments of most regions of the continental United States (with a focus on the Central Plains). The hydrological signatures on the other hand show a clearer pattern. Here, the mean winter discharge, Q95 and the mean annual discharge have a narrow range. This shows that catchments with very different attributes can produce very similar discharge characteristics, as the different attributes seems to cancel each other out in their influence on the discharge.~~

**Cluster 3** is the smallest cluster with only seven catchments. Those are all located in the Northwestern Forested Mountains. Their most distinct feature is their strong negative precipitation seasonality (indicating a strong precipitation peak in the winter) (Figure 6, 7). They also experience high precipitation events mostly in winter falling as snow. Hydrologically, their most distinct features is the very high mean summer discharge and high runoff ratio (Figure 8). This is probably caused by the large amounts of snow melt in late spring and early summer. The catchments of Cluster 3 have the largest overall snow storage with mean maximum value of over 600 mm. Overall, the catchments in this cluster seem to be, from a hydrological point of view, the most extreme in the overall CAMELS data set. This can be seen in their varying discharge patterns. The uniting pattern is their large peak discharge during summer and their extreme values in the PCA space (indicating much higher values for the hydrological signatures in comparison with the other catchments) (Figure 3).~~is the smallest cluster with only seven catchments. Those are all located in the Northwestern Forested Mountains. Their most distinct feature is their uniform high cover with forest. They also experience high precipitation events only seldom and precipitation is snow half of the time. Hydrologically~~

~~their most distinct features is their very high mean summer discharge and high runoff ratio, which is probably caused by the large amounts of snow these catchments receive.~~

**Cluster 4** is, as cluster 3, located in the Northwestern Forested Mountains, with the exception of four catchments that are located in Florida (Figure 4). This cluster is another example of different catchment attributes being able to create similar discharge characteristics concerning the signatures used, while having very different catchment attributes (Figure 6). The catchments have overall low discharge and few high flow events, except one large peak in the mid of the summer, which is caused by melting snow in the northern catchments and strong rainfalls in Florida. Their catchment attributes vary widely, especially in all attributes that are related to elevation (e.g. fraction of precipitation falling as snow) (Figure 7), which is to be expected when some of the catchments are located close to the sea in the southeast, while others are mountainous.~~is also located in the Northwestern Forested Mountains, with the exception of four catchments that are located in Florida. This again is an example of different catchment attributes being able to create similar discharge characteristics concerning their signatures, while having different catchment attributes. The catchments have overall low discharge and few high flow events, while their catchment attributes vary widely, especially in all attributes that are related to elevation (e.g. fraction of precipitation falling as snow).~~

**Cluster 5** includes only few catchments (n=9). They are all located at regions in the northern part of the Marine West Coast Forests (Figure 4). This is the region in the continental US that receives the highest precipitation (>2000 mm year), which is reflected in their discharge characteristics (Figure 6, 8). These catchments have the highest discharge in the whole dataset, especially in the early summer, due to a combination of high precipitation and snowmelt. They also experience only few high precipitation events as they receive large amounts of rain and snow most of the year, with a distinct very high peak in the winter months. The catchments are uniformly covered by almost 100% of forest.~~, has only few catchments (n = 9). They are all located at regions in the northern part of the Marin West Coast Forests. This is the region in the continental US that receives the highest precipitation (> 2000 mm year). This is mirrored in their discharge characteristics. These catchments have the highest discharge in the whole dataset, especially in the summer. They are also uniformly covered by almost 100 % of forest. They also experience only few high precipitation events as they get rain and snow more or less constantly in the same amount.~~

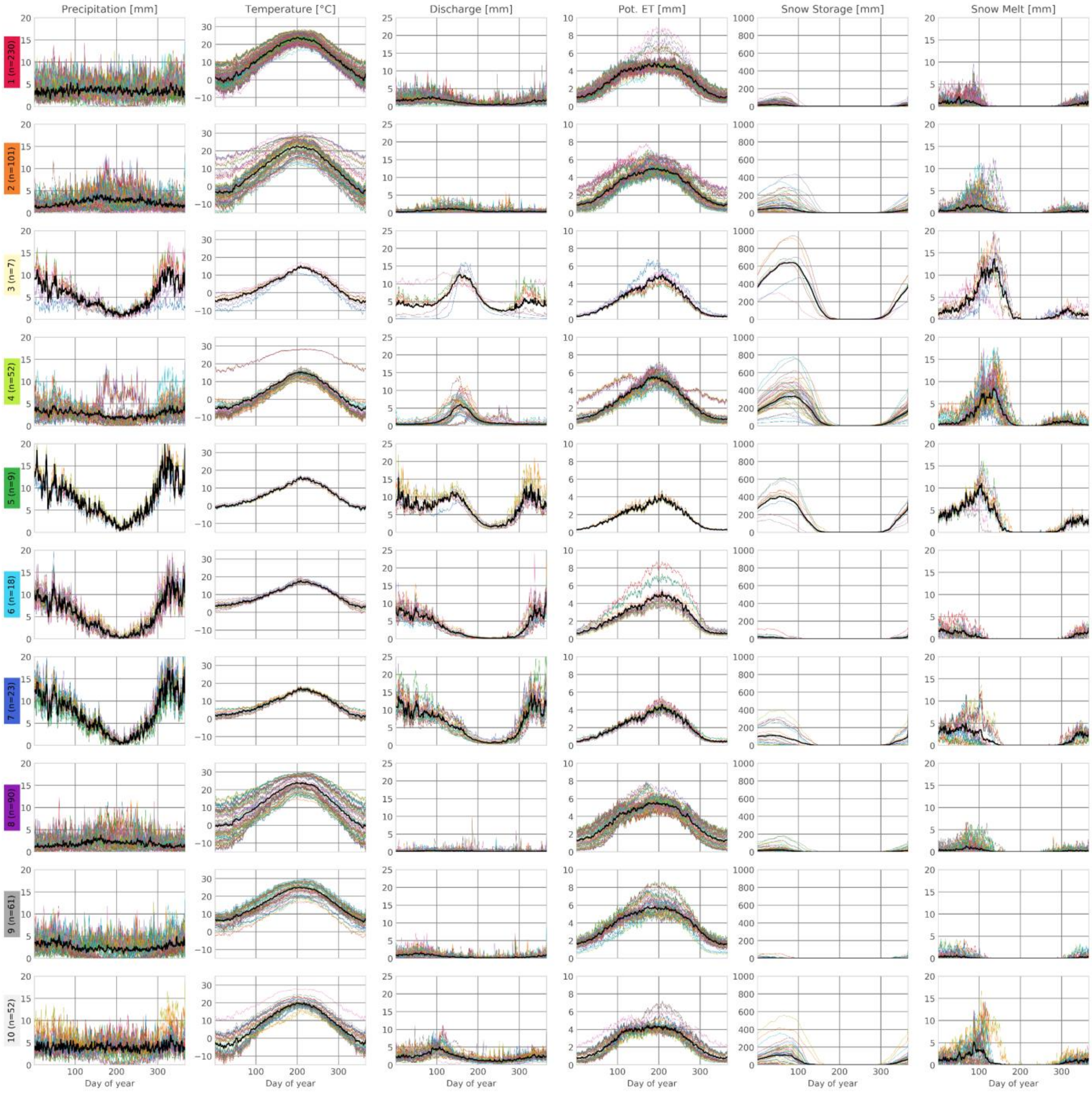



**Figure 6: Meteorological attributes of the clustered CAMELS catchments averaged by day of the year. Potential Evapotranspiration (Pot. ET) was calculated with Hargreaves-Samani (Samani, 2000). Snow storage and melting was calculated using a temperature based approach described (Massmann, 2019). Black lines indicate the mean of all cluster members. Colored lines represent the individual catchments.**

**Cluster 6** is located in the Marine West Coast Forest, but in contrast to Cluster 5, they cover the whole region and not only the northern part (Figure 4). The catchments are very similar in their attributes and discharge characteristics to Cluster 5, with the exception of lower discharges and runoff ratios (Figure 7, 8). This is caused by slightly lower precipitation in comparison with Cluster 5. Cluster 6 experiences the most negative precipitation seasonality across all clusters, with almost all precipitation falling in the winter month. Due to this seasonality and the lower precipitation in the summer, the catchments of this cluster uniformly dry out almost completely in the late summer (Figure 6).~~catchments are also located in the Marine West Coast Forest, but cover the whole region and not only the northern part like Cluster 5. The catchments are very similar in their attributes and discharge characteristics to Cluster 5, with the exception of a lower discharge and runoff ratio. This might be caused by a slightly lower precipitation in comparison with Cluster 5.~~

**Cluster 7** is also located in the same region as Cluster 5 and 6 (Marine West Coast Forests) (Figure 4). Concerning the catchment attributes and the discharge characteristics, it is located between Cluster 5 and 6. So, Cluster 5 to 7 all cover the same region and differ in their mean summer discharge, which is caused by variations in elevation and location (Figure 7). Cluster 7 has higher subsurface permeabilities than cluster 6, which might explain the differences in hydrological behavior, even though the overall attributes of both clusters are rather similar. For example, Cluster 7 has an overall lower discharge than Cluster 5, but does not dry out during the summer as Cluster 6 does (Figure 6). This might be due to the larger amount of snow it receives in comparison with Cluster 6 and its lower evapotranspiration.~~is also located in the same region as Cluster 5 and 6 (Marine West Coast Forests). Concerning the catchment attributes and the discharge characteristics, it is located between Cluster 5 and 6. So, Cluster 5 to 7 all cover the same region and differ in their mean summer discharge, which is caused by slight variations in elevation and location.~~

**Cluster 8** is the most arid cluster (Figure 7). All of the catchments are located in western parts of the Great Plains and in the North American Deserts (Figure 4). They are characterized by an overall low water availability and high evaporation, which is shown in the very low mean annual discharge and runoff ratio (Figure 6, 8). This also results in low values for the LAI. However, the frequency of high precipitation events is high.~~is the overall most arid cluster catchments. All of the catchments are located in western parts of the Great Plains and in the North American Deserts. They are shaped by an overall little availability of water and high evaporation, which is shown in the very low mean annual discharge and runoff ratio. This also results in low values for the LAI. However, the frequency of high precipitation events is high.~~

**Cluster 9** covers all southern states of the United States (Figure 4). The catchments here are quite similar to Cluster 8, but show a lower precipitation seasonality and a higher forest cover and green vegetation (Figure 7). In addition, all catchments of this cluster are in relative close proximity to the sea. The uniting factor in this cluster seems to be the very low snow fraction and the high evapotranspiration (Figure 6, 7).~~covers all southern states of the United States. The catchments here are quite similar to Cluster 8, but show a lower seasonality (as indicated by their lower half flow date) and a higher forest cover and green vegetation.~~

**Cluster 10** catchments are all located in the Appalachian Mountains (Figure 4). The mean elevation is higher than of most other clusters and the catchments also depict a low aridity and a very high forest cover (Figure 7). Their discharge characteristics are similar to that of the Marine West Coast Forests Clusters 5 to 7 (Figure 6, 8). However, they receive less water than those catchments. Cluster 10 covers the same ecoregion as Cluster 1, but has a distinct behavior due to its mountainous character, which can be seen in the higher seasonality of the discharge. This is probably caused by the larger snow cover, with a returning snow melt discharge peak in spring.~~catchments are located in the Appalachian Mountains. The mean elevation higher than most other clusters and the catchments also have low aridity and a very high forest cover. Their discharge characteristics is similar to the Marine West Coast Forests of Cluster 5 to 7. However, they receive less water than those catchments and experience a higher seasonality (as indicated by the higher mean half-flow date).~~

Overall, we can see similar trends for some of the cluster. We identified four distinct groups. The general similarities of the clusters are also represented by their distance and position in the PCA space (Figure 3).

- Group 1 (Cluster 1, 2, 8, 9): low seasonality in precipitation and discharge; located in the eastern US; due to low slope inclinations, water takes a long time to reach the outlet.
- Group 2 (Cluster 3, 4): dominant summer peak of discharge caused by rapid snow melt; mostly located in the mountains of the western US; differ in precipitation inputs.
- Group 3 (Cluster 5, 6, 7): located in the Northwestern Forested Mountains; characterized by high precipitation amount and seasonality, but more or less extreme versions.
- Group 4 (Cluster 10): located in the Appalachian mountains; share characteristics with Group 1, though influenced by higher elevations and steeper slopes.

Those groups of clusters are similar to the ones found by (Berghuijs et al., 2014), even though they used a very different method to derive them. The main difference in the groups is probably caused by how we structure the clusters and groups in the eastern US, due our clusters being more influenced by the Appalachian Mountains. However, both approaches deliver similar results overall.

The question remains: what is the right numbers of clusters? Though even we did find four distinct groups, having only four clusters would probably be too little, as the clusters in the groups show a wide range of behavior (Figure 3, 7, 8, Table 2). There are catchment attributes, which we did not take into account, but which could further split up the clusters (e.g. the shape of the catchments). However, this study considered the catchment attributes that are usually considered as being important. The fact that the clusters contain different numbers of catchments can be explained by their distances in the PCA space (Figure 3). Many of the catchments are rather similar. This produces some clusters with most of the catchments. However, we also have some extreme catchments (e.g. Cluster 3 and 5), which are very different to the bulk of the catchments in the CAMELS dataset. Thus, even though some of our presented clusters are quiet small in number, they are needed to capture their extreme hydrological behavior. Our results show that some of the clusters follow the boundaries of the ecoregions in the US very directly (Cluster 1), while others do not (Cluster 9). The worlds of ecology and hydrology are sometimes shaped by the same forcing, but not always.

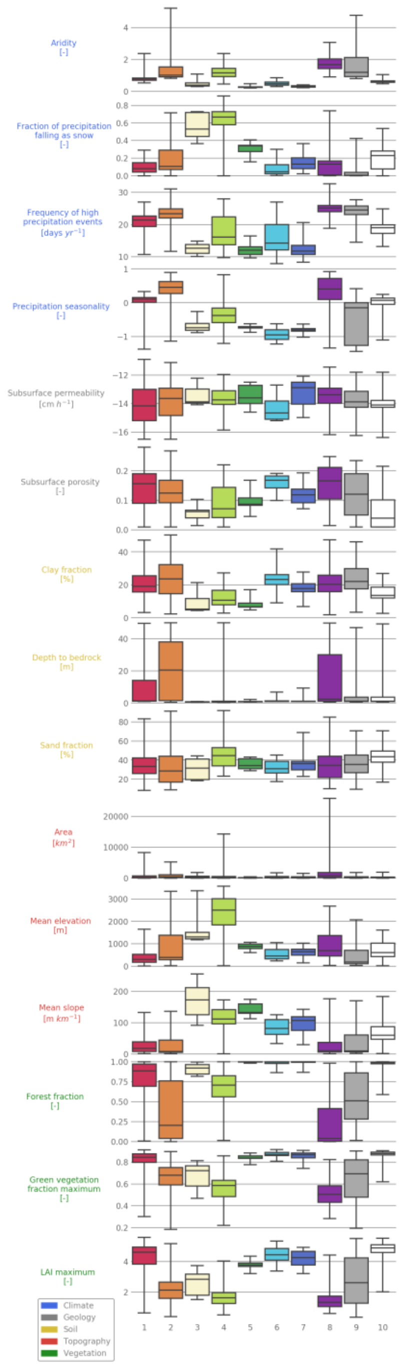



**Figure 7: Boxplots of the catchment attributes of the clusters**

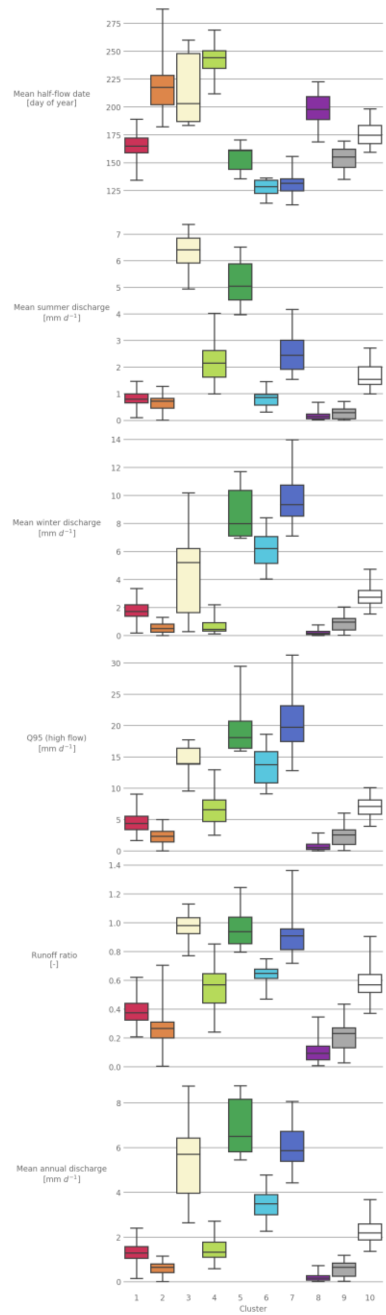



Figure 8: Boxplots of the hydrological signatures of the clusters.

**Table 2: Properties of the catchment clusters. Typical signatures/attributes refers to the signature/attribute of the cluster with the lower coefficient of variation scaled by the mean coefficient of variation of the whole dataset. Dominating attribute refers to the catchment attribute that has the highest weighted $R^2$.**

| Cluster | n | Main Region | Typical signature | Typical attribute and their manifestation | Dominating attribute |
|---|---|---|---|---|---|
| 1 | 230 | Southeastern and Central Plains | Low mean winter discharge | Low aridity | Aridity |
| 2 | 101 | Central Plains (with scattered catchments all over western US) | High mean half-flow date | ~~Mid to low depth to bedrock~~High precipitation seasonality | Green vegetation fraction maximum |
| 3 | 7 | Northwestern Forested Mountains | High mean summer discharge | ~~High forest fraction~~Low precipitation seasonality | Fraction of precipitation falling as snow |
| 4 | 52 | Northwestern Forested Mountains and Florida | High mean half-flow date | Mid frequency of high precipitation events | ~~Fraction of precipitation falling as snow~~Precipitation seasonality |
| 5 | 9 | Northern Marine West Coast Forests | High mean summer discharge | Very high forest fraction | Forest fraction |
| 6 | 18 | Marine West Coast Forests | Mid runoff ratio | ~~Very high forest fraction~~Low precipitation seasonality | Aridity |
| 7 | 23 | Western Cordillera (Part of Marine West Coast Forests) | High mean winter discharge | ~~Very high forest fraction~~Low precipitation seasonality | Fraction of precipitation falling as snow |
| 8 | 90 | Great Plains and North American Deserts | Mid mean half-flow date | High frequency of high precipitation events | ~~Aridity~~Precipitation Seasonality |
| 9 | 61 | All southernmost states of the US | Low mean half-flow date | High frequency of high precipitation events | Aridity |
| 10 | 52 | Appalachian Mountains | Low mean winter discharge | High forest fraction | Mean elevation |

~~The derived importance of the catchment attributes in the clusters is highly variable and partly differs from the order of importance in the overall dataset (compare Figure 1 and Figure 5). For Cluster 1 (Southeastern and Central Plains), 6 (Marine West Coast Forests), 8 (Great Plains and Deserts) and 9 (Southern states) aridity has the clearest connection to the clusters. However, this is not the case for the remaining catchment clusters. For Cluster 3 (Northwestern Forested Mountains), 4 (Northwestern Forested Mountains and Florida) and 7 (Western Cordillera) the clearest connection is to the fraction of precipitation falling as snow. However, for Cluster 3, and 4 many other catchment attributes have a weighted $R^2$, which is almost as high as the one for the fraction of precipitation falling as snow.~~

**3.5 Importance of the catchment attributes in the clusters**

The individual importance of the catchment attributes in the clusters is variable and partly deviates from the order of importance in the overall dataset (compare Figure 2 and Figure 9). For Cluster 1 (Southeastern and Central Plains), 6 (Marine West Coast Forests) and 9 (coastal states) aridity has the highest weighted coefficient of determination in the clusters. For Cluster 3 (Northwestern Forested Mountains) and 7 (Western Cordillera) the highest relevance is found for the fraction of precipitation falling as snow. For the remaining clusters it is precipitation seasonality (Cluster 4 (Northwestern Forested Mountains), Cluster 8 (Great Plains and Deserts)), the green vegetation fraction maximum (Cluster 2 (Central Plains)) and the mean elevation (Cluster 10 (Appalachian Mountains)). We can also see that some clusters have one dominating catchment attribute (investigated by the coefficient of determination e.g. aridity in Cluster 1, compare Figure 9), while for other clusters, all attributes seem equally important (e.g. Cluster 8). Overall, the western clusters (west of the 100th meridian) display the highest weighted $R^2$ with:~~The derived importance of the catchment attributes in the clusters is variable and differs from the order of importance in the overall dataset (compare Figure 1 and Figure 5). For Cluster 1 (Southeastern and Central Plains), 6 (Marine West Coast Forests), 8 (Great Plains and Deserts) and 9 (Southern states) aridity has the clearest connection to the clusters. However, this is not the case for the remaining catchment clusters. For Cluster 3 (Northwestern Forested Mountains), 4 (Northwestern Forested Mountains and Florida) and 7 (Western Cordillera) the clearest connection is to the fraction of precipitation falling as snow. However, for Cluster 3, and 4 many other catchment attributes have a weighted $R^2$, which is almost as high as the one for the fraction of precipitation falling as snow. In addition, all catchment attributes have a high weighted $R^2$ in Cluster 3, while the weighted $R^2$ is low for all catchment attributes in Cluster 4. For the remaining clusters, it is green vegetation maximum (Cluster 2, Central Plains), forest fraction (Cluster 5, Northern Marine West Coast Forest) and mean elevation (Cluster 10, Appalachian Mountains). Overall, the western clusters (west of the 100th meridian) have the highest weighted $R^2$ with the:~~

- Fraction of precipitation falling as snow (Cluster 3, 7)
- Precipitation seasonality (Cluster 4)
- Forest fraction (Cluster 5)

- Aridity (Cluster 6)

eastern clusters (east of the 100th meridian) with:

- Aridity (Cluster 1)
- Mean elevation (Cluster 10)

clusters equally present in west and east with:

- Green vegetation fraction maximum (Cluster 2)
- Aridity (Cluster 9)
- Precipitation seasonality (Cluster 8)
- ~~Fraction of precipitation falling as snow (Cluster 3, 4, 7)~~
- ~~Forest fraction (Cluster 5)~~
- ~~Aridity (Cluster 6)~~

~~The eastern clusters (east of the 100th meridian) with the:~~

- ~~Aridity (Cluster 1)~~
- ~~Mean elevation (Cluster 10)~~

~~The clusters equally present in west and east with the:~~

- ~~Green vegetation fraction maximum (Cluster 2)~~
- ~~Aridity (Cluster 8, 9)~~

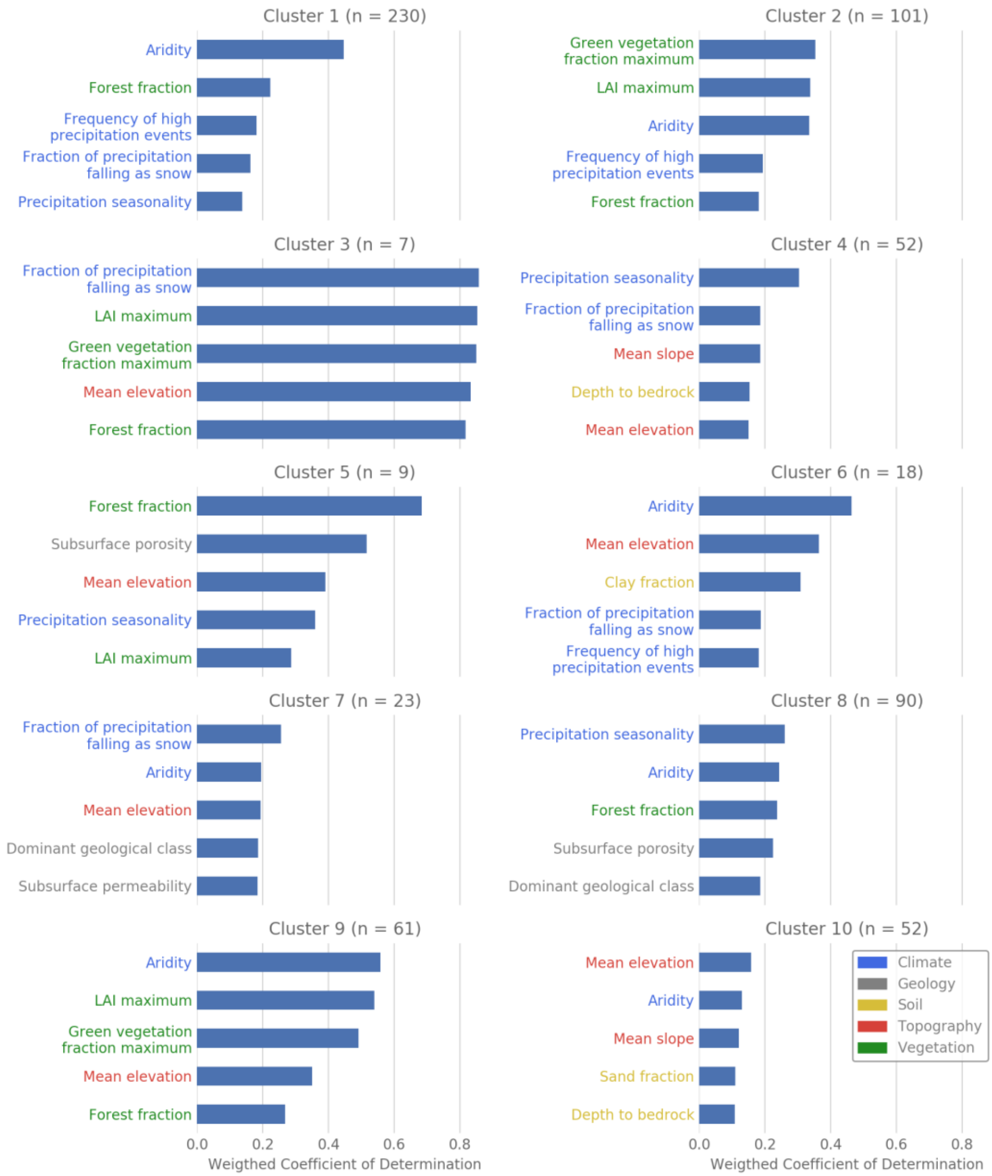



**Figure 9~~5~~: Importance of the catchment attributes evaluated by the quadratic regression~~. F~~for the catchment clusters. Attributes colored according to their catchment attribute class.**

Keeping the correlation coefficients displayed in Figure 1 in mind, we see that climate is the most important factor in almost all clusters, as the vegetation attributes are highly correlated with the climate attributes. The only exception is Cluster 10 in which mean elevation is the most important catchment attribute. However, the catchment attributes in Cluster 10 have overall low $R^2$ values and the mean elevation is directly followed by the aridity. This again shows that climate seems to be the dominating factor for catchment behavior, as found in other large sample studies (e.g. (Berghuijs et al., 2014; Kuentz et al., 2017)). Nevertheless, if one takes a closer look at the data set, more detailed, regional correlations with regard to individual climate variables can be determined. For example, Cluster 1 is defined by the aridity, while Cluster 4 seems to be much more influenced by the precipitation ~~Overall~~seasonality. Overall, it is feasible to link dominating catchment attributes to the hydrological behavior. While it is straightforward in some regions of the US, it is more challenging in others. We link this to the signal of the climatic forcing being more superimposed by other catchment attributes, which results in a less clear connection between its hydrological behavior and the climate. This hints that climate and catchment attributes are more intertwined in those areas and indicates regions where different types of hydrological runoff generation processes are existing. Furthermore, it indicates regions where hydrological predictions in ungauged basins~~, we found that it is relatively easy to link the dominating catchment attributes to the hydrological behavior, in some regions of the US. However, it is more challenging in others. We link this to the signal of the climatic forcing being more influenced by other catchment attributes, which results in a less clear connection between hydrological behavior and climate. This hints that climate and catchment attributes are more intertwined in those areas and indicates regions where different types of hydrological model structures are needed. Furthermore, it indicates regions where hydrological predictions in ungauged basins~~ (Hrachowitz et al., 2013) can become very challenging, as the interplay of the available meteorological ~~-~~data and catchment-attributes ~~data~~ cannot sufficiently explain the hydrological characteristics. Those findings also highlight one current discrepancy between large sample and single catchment studies. While large sample studies, especially the very large ones, identify climate as being most important for the hydrological behavior (e.g. (Addor et al., 2018; Kuentz et al., 2017)), smaller studies (e.g. (Chiverton et al., 2015; Pfister et al., 2017)) and single catchment studies (e.g. (Floriancic et al., 2018) often identifiy the geology or soils as being very important. This might be linked to the overall problem of scales in hydrology as different scales of soil/geology and climate have different effects and different data accuracy ~~(Addor et al. 2018)~~(Addor et al., 2018; Blöschl, 2001). In addition to this, the overall scale might also come into play. Smaller sample studies often ~~have~~ compare catchments that are not far away from each other and probably ~~receive~~ have similar climate forcings. Thus, the differences in hydrological behavior can only be caused by catchment attributes other than climate. Therefore, larger and smaller sample studies might be looking at different things. While very large sample studies capture what drives catchments on large scales (the climate), smaller studies look at how this climatic signal is transferred to discharge by the catchment attributes.

~~In the next step, we linked the abovementioned findings to the differences between the correlations of the catchment attributes with each other in the eastern and western parts of the continental US (Figure 6). While aridity is the most important catchment attribute, when looking at all catchments at the same time (Figure 1), this does not hold true for most of the single clusters (Figure 5). Yet, the factors with the highest weighted coefficient of correlation might simply be proxies for aridity. To test this, we scrutinized the correlation between the catchment attributes with each other, separated by East and West (Figure 6). The~~

~~western US (Figure 6a) and eastern US (Figure 6b) show high differences in the way the catchment attributes correlate with each other (Figure 6c). The main differences are in the mean elevation, the fraction of precipitation falling as snow, and the LAI maximum. For example, in the western US the mean elevation has a high correlation (r = 0.8) with the fraction of precipitation falling as snow. In the eastern US however, this correlation is much smaller (r = 0.4). This is probably caused by the overall higher elevation in the western US. In addition, in the western US, the fraction of the precipitation falling as snow does not correlate with the aridity (r = 0.1), while the forest fraction does (r = -0.8). Thus, the forest fraction is linked very directly to the climate in this region. Therefore, aridity (and the highly correlated forest fraction) have the highest weighted R² in two out of the five clusters in the western US. Only two clusters are mostly located in the eastern US (Cluster 1 and 10). Here, aridity and the mean elevation have the highest weighted R² with the hydrological behavior. The mean elevation has a medium correlation with the aridity. Hence, the hydrological behavior in the eastern US is most highly correlated with aridity, which is not the case for the western US. There, the fraction of precipitation falling as snow is more prevalent. Those results imply that aridity is a good indicator for the discharge characteristics in the eastern US and only mediocre in the West.~~

**~~Table 2: Properties of the catchment clusters. Typical signatures/attributes refers to the signature/attribute of the cluster with the lower coefficient of variation scaled by the mean coefficient of variation of the whole dataset. Dominating attribute refers to the catchment attribute that has the highest weighted R².~~**

| ~~Cluster~~ | ~~n~~ | ~~Main Region~~ | ~~Typical signature~~ | ~~Typical attribute and their manifestation~~ | ~~Dominating attribute~~ |
|---|---|---|---|---|---|
| ~~1~~ | ~~230~~ | ~~Southeastern and Central Plains~~ | ~~Low mean winter discharge~~ | ~~Low aridity~~ | ~~Aridity~~ |
| ~~2~~ | ~~101~~ | ~~Central Plains (with scattered catchments all over western US)~~ | ~~High mean half-flow date~~ | ~~Mid to low depth to bedrock~~ | ~~Green vegetation fraction maximum~~ |
| ~~3~~ | ~~7~~ | ~~Northwestern Forested Mountains~~ | ~~High mean summer discharge~~ | ~~High forest fraction~~ | ~~Fraction of precipitation falling as snow~~ |
| ~~4~~ | ~~52~~ | ~~Northwestern Forested Mountains and Florida~~ | ~~High mean half-flow date~~ | ~~Mid frequency of high precipitation events~~ | ~~Fraction of precipitation falling as snow~~ |
| ~~5~~ | ~~9~~ | ~~Northern Marine West Coast Forests~~ | ~~High mean summer discharge~~ | ~~Very high forest fraction~~ | ~~Forest fraction~~ |
| ~~6~~ | ~~18~~ | ~~Marine West Coast Forests~~ | ~~Mid runoff ratio~~ | ~~Very high forest fraction~~ | ~~Aridity~~ |

| | | | | | |
|---|---|---|---|---|---|
| ~~7~~ | ~~23~~ | ~~Western Cordillera (Part of Marin West Coast Forests)~~ | ~~High mean winter discharge~~ | ~~Very high forest fraction~~ | ~~Fraction of precipitation falling as snow~~ |
| ~~8~~ | ~~90~~ | ~~Great Plains and North American Deserts~~ | ~~Mid mean half-flow date~~ | ~~High frequency of high precipitation events~~ | ~~Aridity~~ |
| ~~9~~ | ~~61~~ | ~~All southernmost states of the US~~ | ~~Low mean half-flow date~~ | ~~High frequency of high precipitation events~~ | ~~Aridity~~ |
| ~~10~~ | ~~52~~ | ~~Appalachian Mountains~~ | ~~Low mean winter discharge~~ | ~~High forest fraction~~ | ~~Mean elevation~~ |

### 3.~~4~~6 Differences in clusters in comparison with other hydrological clustering studies

Compared to the clustering results of Kuentz et al. (2017), who derived their cluster from European catchments by an analogous method, some similarities can be found. Like them, this study here also found one cluster (Cluster 2) that does not have any distinct character. However, only around one sixth of the CAMELS catchments belongs to this Cluster 2, while Kuentz et al. (2017) had one third of their catchments in a cluster without distinct features. Therefore, our selection of hydrological signatures seems to allow a better identification of hydrological similarities. However, all catchments in CAMELS are mostly without human impact (Addor et al., 2017), while many catchments in the study of Kuentz et al. (2017) are under human influence. ~~This influence might overlay potentially apparent patterns~~This human influence might mask otherwise apparent patterns. Kuentz et al. (2017) also found two clusters that contain mostly mountainous catchments. These show a similar behavior to Cluster 3 (Northwestern Forested Mountains) and Cluster 10 (Appalachian Mountains) found in Figure ~~3~~4. The main difference between their findings and this study here is Cluster 8, as it contains very arid catchments (with some being located in deserts). Obviously, this cluster cannot be found in Europe as Europe has no real deserts. Still, there is some similarity with their cluster of Mediterranean catchments as both are dominated by aridity. Summarizing, in their study and this study catchments are mainly clustered in groups of desert/arid catchments, mountainous catchments, mid height mountains with high forest ~~shares~~fraction, wet lowland catchments and one cluster of catchments that do not show a very distinct behavior and therefore do not fit in the other clusters (Table 2). One possible explanation for this unspecific behavior might that many catchments have one or two important attributes that dictate most of their behavior, but which are different from other cluster members. For example, desert catchments are relatively easy to identify, as they are dominated by heat and little precipitation.

A European upland catchment on the other hand ~~have~~ has several more influences such as snow in the winter, heat in the summer, varying land use and strong impact of seasonality. Here, many influences overlap each other and make it thus difficult to identify a single causes, see also the discussion by Trancoso et al. (2017) that goes in a similar direction. Those overlapping influences are probably also the reason why catchment classification studies often find clusters where one or two clusters that include a large number of catchments, while most other cluster only contain few catchments (Coopersmith et al., 2012; Kuentz et al., 2017). Therefore, it is quite difficult to confirm the 'wish' of the hydrological community to have homogenous catchment groups with only a few outliers (e.g. (Burn, 1997)), because catchments are complex systems with a high level of self–organization arising from co-evolution of climate and landscape properties, including vegetation (Coopersmith et al., 2012). Accordingly, it requires many separate clusters to separate those multi-influence catchments into homogenous groups. Still, the cluster found here might capture much of the variety present in the United States, as they roughly follow ecological regions (McMahon et al., 2001), which has been stated as a hint of a good classification (Berghuijs et al., 2014). In addition, this study shows that using clusters derived from principal components of hydrological signatures create meaningful groups of catchments with similar attributes (Figure ~~A2, A3~~6, 7, 8). Those clusters also show distinct spatial patterns (Figure 4~~3~~). Similar results were also found in other studies that used the same method (Kuentz et al., 2017; McManamay et al., 2014), but based them on partly different hydrological signatures. Therefore, the principal components of hydrological signatures can be used as a measure of similarity between catchments. They represent the "essence" of all hydrological signatures used. Our results also show that it is difficult to link those catchment clusters to simple averaged measures of catchment attributes. While some clusters have very clear connections to the attributes, others have no catchment attribute that could easily explain the behavior of the catchments. This hints, that some catchments are easier to explain (in a hydrological sense) than others. Those difficulties might be an artifact of the averaged catchment attributes or be caused by complex catchment reaction, forced by intertwined climate and catchment attributes. Which in turn, might indicate an equifinality of catchment response.

### 3.~~5~~7 Comparing catchment clusters based on hydrological behavior and climate

Besides hydrological behavior, climate is often used to sort catchments into similar groups (e.g. Berghuijs et al., 2014; Knoben et al., 2018). Therefore, we are interested if both approaches deliver comparable results. To evaluate this, we contrasted our results to the commonly used Koeppen-Geiger climate classification (Beck et al., 2018) (Figure ~~7~~10) and recently published approach of Knoben et al. (2018), who sorted climate along three continuous axis of aridity, seasonality and fraction of precipitation falling as snow (Figure 11~~8~~). The resulting clusters based on climate and hydrology should be the same, if climate is the dominating driver of hydrological behavior in every catchment. Yet, this is not the case for the Koeppen-Geiger classification. In every hydrological cluster are at least two different climates regarding the Koeppen-Geiger classification, ranging up to eight different climatic regions for Cluster 2 and 8 (those even include deserts and very cold regions). Thus, the

Koeppen-Geiger classification seems unable to capture the essential drivers of hydrological behavior. A critique also raised in other studies (e.g. Haines et al. (1988); Knoben et al. (2018)).

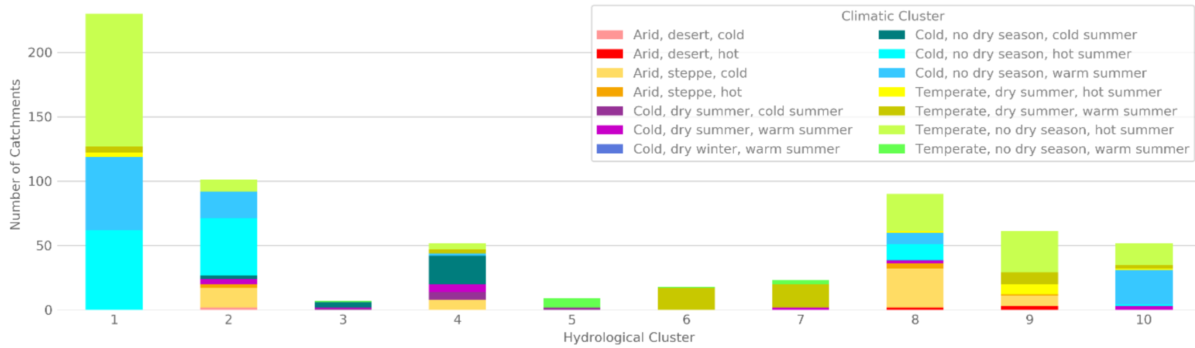



**Figure ~~7~~10: Membership of Koeppen-Geiger clusters (Beck et al. (2018)) in the hydrological clusters.**

The picture is less clear concerning the climatic index space of Knoben et al. (2018) (Figure 11~~8~~a). Due to the continuous nature of the approach of Knoben et al. (2018), there are no clear boundaries as in the Koeppen-Geiger classification. Still, there are some emerging patterns. For example, according to the approach of Knoben et al. (2018) Cluster 1 is mainly defined by a relatively arid climate, with some seasonal variability and little to no snow. This is in line with our analysis of the most influential catchment attributes for this cluster, as we identified aridity as the main driver. ~~Contrastingly, clusters where we could not identify a clear dominating catchment attribute, e.g. Cluster 4 (located in the Northwestern Forested Mountains and Florida) (Figure 5), also have no clear clustering in the climate index space. The catchments of those clusters can be found in the space of the climatic indices of Knoben et al. (2018) with very different aridity, seasonality and fraction of the precipitation falling as snow.~~ There seem to be regions were the forcing signal of the climate is transferred more directly to a streamflow response than in others. However, this does not mean that climate is unimportant in those regions. Either the climate forcing signal is changed more through other attributes of the catchment, or the mean values describing the climate do not properly reflect the variability of the climate in the single catchments. This leads to less clear correlation between the climate and the hydrological behavior. Interestingly, when we look at the single hydrological signatures in the climate index space (Figure 11~~8~~b, A~~4~~2) we see a very clear connection between the single hydrological signatures and the climate. This direct connection of the signatures used was also found by Addor et al. (2018). Our results and the comparison show that the complex hydrological behavior, captured in a range of hydrological signatures, does not simply follow the climate only, even though the individual signatures do. Still, all signatures combined seem to capture a dynamic, which is climatic in origin, but is shaped through the attributes of the catchments (like vegetation and soils (Berghuijs et al., 2014)). Therefore, to find truly similar catchments, using climate characteristics only, is probably not sufficient (see also (Addor et al., 2018; Knoben et al., 2018; Kuentz et al., 2017)).

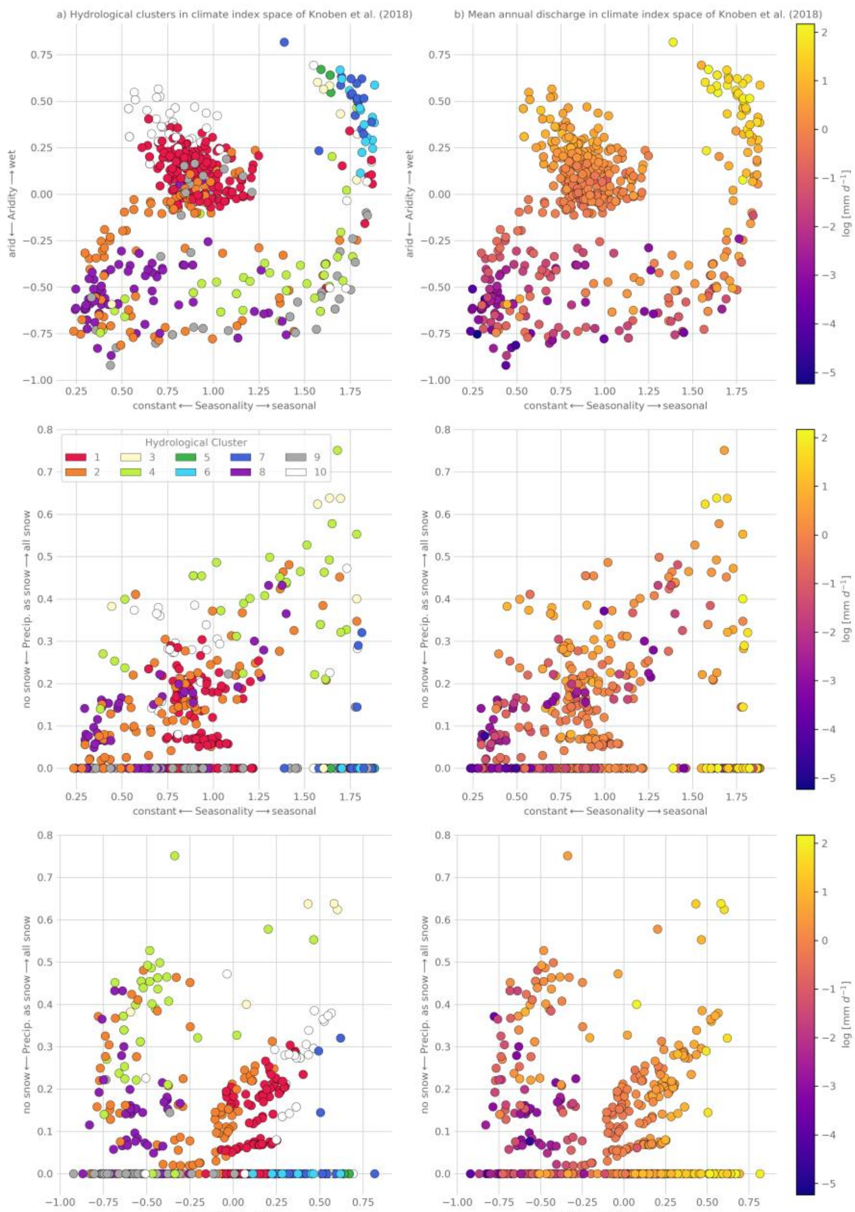



**Figure 11~~8~~: a) Comparison of the hydrological clustering of this study with the climate index space of Knoben et al. (2018). Single dots show the catchments and are colored by their hydrological clusters. b) Mean annual discharge for all catchments in the climate index space of Knoben et al. (2018). Single dots show the catchments and are colored according to the value of the mean annual discharge. The log of the mean annual discharge is used to show the relative differences between the catchments. For a depiction of all hydrological signatures used, see Figure A~~4~~2.**

## 4 Summary and conclusion

This study explored differences in the catchment characteristics between the eastern and western US, the properties and location of catchment clusters based on hydrological signatures, the importance of catchment attributes for those clusters and how this study relates to other clustering studies and methods. We found that the correlations catchment characteristics are quite similar for the eastern and western US with the exception of mean elevation, snow, geology and the leaf area index. For the overall CAMELS data set climate seems to be the most important factor for the hydrological behavior. However, depending on the location either aridity, snow or seasonality were most important. The clusters derived from the hydrological signatures partly follow the eco regions in the US and can combined into four groups of general behavior trends. Still, similar catchments can be quite far away from each other. We also found that most of the catchments have a rather similar discharge behavior, while only some, more extreme catchments, derivate from that main trend. This might be a hint why it is so difficult to clusters catchments, as those single extreme catchments are quite unique and do not fit together well with other catchments. ~~This study explored the influence of catchment attributes on the discharge characteristics in the CAMELS dataset. We found that over the whole dataset climate (especially aridity) is the most important factor for the discharge characteristics. This changes when we take a closer look at clusters that are derived from specific hydrological signatures. For the clusters in the eastern US, aridity is still the most important catchment attribute. In the western US however, the amount of snow is more important. In addition, in the western catchments the hydrological behavior is less correlated with the remaining catchment attributes. It seems like the clear climatic signal in the east is dampened in the west. This might be caused by a higher influence of other catchment attributes like elevation and vegetation. A similar effect can be found, when we compare how catchments align along hydrological and climatic axes. While some hydrological clusters align along a relatively narrow range of values of the climatic indices, others are found in very contrasting climates. Summarizing~~We also found, that there are differences of how directly the signal of forcing climate can be found again in the hydrological behavior. This explains why catchments often show a surprisingly similar behavior across many different climate and landscape properties (Troch et al., 2013) and why the most hydrologically similar catchment can be hundreds of kilometers away. Those findings also relate to the paradox that small scales/single catchment studies identify geology/soils as most important for the hydrological behavior, while large sample studies usually find the climate as most important. This might simply be influenced by spatial proximity. Small scale studies look at catchment which all have a similar climatic forcing and thus only the other catchment attributes can be the cause of differences in hydrological behavior. Large sample studies on the other hand consider catchment from a wider area and thus attribute the differences in behavior to climate.

The aggregated data used in this study might level out the variability of the catchment attributes in the single catchment, but it also indicates that there is a kind of equifinality in the behavior of catchments. Different sets of intertwined climate forcing and catchment attributes could lead to a very similar overall behavior, not unlike to hydrological models that produce the same discharge with different sets of parameters.

We acknowledge that the results are dependent on the amount and size of the clusters, the catchment attributes considered and the hydrological signatures used. Still, we think that the CAMELS dataset offers an excellent overview of different kinds of catchments in contrasting climatic and topographic regions. In addition, this study shows that using hydrological signatures with high spatial predictability results in hydrological meaningful clusters, which show consistent low flow behavior, even though those low flows were not explicitly considered. However, it seems that even a comprehensive dataset like CAMELS, does not allow an easy way to find a conclusive set of clusters for catchments. For future research, we recmonned~~it might be a worthwhile pathway~~ to include measures of spatial variability of the climate in the single catchments and to look into the single clusters in more depth. This might help to prove, if a less clear climatic signal is caused by intra-catchment variability of the climate or a larger influence of other catchment attributes.

**Data availability**

The CAMELS dataset can be found at https://ncar.github.io/hydrology/datasets/CAMELS_timeseries and is described in Addor et al. (2017). The cluster numbers together with the CAMELS catchment ID and the climatic indices can be found in the supplement of this paper.

**Code availability**

The code used for this study can be found at Jehn (2018).

**Author contribution**

FUJ, LB, TH and PK conceived and designed the study. FUJ did the data analysis. All authors aided in the interpretation and discussion of the results and the writing of the manuscript.

**Competing interests**

The authors declare that they have no conflict of interest.

**Acknowledgment**

We would like to thank Ina Pohle, Marc Vis, Jan Seibert, Wouter Knoben, Andrew Newman and ~~one~~ two anonymous ~~referee~~ reviewers for giving valuable and important feedback in the creation of this paper. We would also like to thank all the people who helped create the CAMELS dataset. Thank you for your work! We further would like to thank the DFG for generously funding the project HO 6420/1-1.

**References**

Addor, N., Newman, A. J., Mizukami, N. and Clark, M. P.: The CAMELS data set: catchment attributes and meteorology for large-sample studies, Hydrology and Earth System Sciences, 21(10), 5293–5313, doi:10.5194/hess-21-5293-2017, 2017.

Addor, N., Nearing, G., Prieto, C., Newman, A. J., Le Vine, N. and Clark, M. P.: A ranking of hydrological signatures based on their predictability in space, Water Resources Research, doi:10.1029/2018WR022606, 2018.

Ali, G., Tetzlaff, D., Soulsby, C., McDonnell, J. J. and Capell, R.: A comparison of similarity indices for catchment classification using a cross-regional dataset, Advances in Water Resources, 40, 11–22, doi:10.1016/j.advwatres.2012.01.008, 2012.

Alvarez-Garreton, C., Mendoza, P. A., Boisier, J. P., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., Lara, A., Puelma, C., Cortes, G., Garreaud, R., McPhee, J. and Ayala, A.: The CAMELS-CL dataset: catchment attributes and meteorology for large sample studies – Chile dataset, Hydrology and Earth System Sciences, 22(11), 5817–5846, doi:10.5194/hess-22-5817-2018, 2018.

Berghuijs, W. R., Sivapalan, M., Woods, R. A. and Savenije, H. H. G.: Patterns of similarity of seasonal water balances: A window into streamflow variability over a range of time scales, Water Resources Research, 50(7), 5638–5661, doi:10.1002/2014WR015692, 2014.

Beven, K. J.: Uniqueness of place and process representations in hydrological modeling, Hydrology and Earth System Sciences, 4(2), 203–213, 2000.

Blöschl, G.: Scaling in hydrology, Hydrol. Process., 15(4), 709–711, doi:10.1002/hyp.432, 2001.

Burn, D. H.: Catchment similarity for regional flood frequency analysis using seasonality measures, Journal of Hydrology, 202(1–4), 212–230, doi:10.1016/S0022-1694(97)00068-1, 1997.

Chiverton, A., Hannaford, J., Holman, I., Corstanje, R., Prudhomme, C., Bloomfield, J. and Hess, T. M.: Which catchment characteristics control the temporal dependence structure of daily river flows?: TEMPORAL DEPENDENCE IN RIVER FLOW, Hydrological Processes, 29(6), 1353–1369, doi:10.1002/hyp.10252, 2015.

Coopersmith, E., Yaeger, M. A., Ye, S., Cheng, L. and Sivapalan, M.: Exploring the physical controls of regional patterns of flow duration curves - Part 3: A catchment classification system based on regime curve indicators, Hydrology and Earth System Sciences, 16(11), 4467–4482, doi:10.5194/hess-16-4467-2012, 2012.

Costanza, R., Wainger, L., Folke, C. and Mäler, K.-G.: Modeling Complex Ecological Economic Systems: Toward an Evolutionary, Dynamic Understanding of People and Nature, in Ecosystem Management, pp. 148–163, Springer New York, New York, NY., 1993.

Floriancic, M. G., Meerveld, I., Smoorenburg, M., Margreth, M., Naef, F., Kirchner, J. W. and Molnar, P.: Spatio-temporal variability in contributions to low flows in the high Alpine Poschiavino catchment, Hydrological Processes, 32(26), 3938–3953, doi:10.1002/hyp.13302, 2018.

Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M. and Andréassian, V.: Large-sample hydrology: a need to balance depth with breadth, Hydrology and Earth System Sciences, 18(2), 463–477, doi:10.5194/hess-18-463-2014, 2014.

Jehn, F. U.: zutn/Catchment-Classification: Release for Zenodo, , doi:10.5281/zenodo.2203639, 2018.

Knoben, W. J. M., Woods, R. A. and Freer, J. E.: A Quantitative Hydrological Climate Classification Evaluated With Independent Streamflow Data, Water Resources Research, 54(7), 5088–5109, doi:10.1029/2018WR022913, 2018.

Kuentz, A., Arheimer, B., Hundecha, Y. and Wagener, T.: Understanding hydrologic variability across Europe through catchment classification, Hydrology and Earth System Sciences, 21(6), 2863–2879, doi:10.5194/hess-21-2863-2017, 2017.

Li, C., Cao, J., Nie, S.-P., Zhu, K.-X., Xiong, T. and Xie, M.-Y.: Serum metabolomics analysis for biomarker of Lactobacillus plantarum NCU116 on hyperlipidaemic rat model feed by high fat diet, Journal of Functional Foods, 42, 171–176, doi:10.1016/j.jff.2017.12.036, 2018.

Massmann, C.: Modelling Snowmelt in Ungauged Catchments, Water, 11(2), 301, doi:10.3390/w11020301, 2019.

McGlynn, B. L. and Seibert, J.: Distributed assessment of contributing area and riparian buffering along stream networks: TECHNICAL NOTE, Water Resources Research, 39(4), doi:10.1029/2002WR001521, 2003.

McMahon, G., Gregonis, S. M., Waltman, S. W., Omernik, J. M., Thorson, T. D., Freeouf, J. A., Rorick, A. H. and Keys, J. E.: Developing a Spatial Framework of Common Ecological Regions for the Conterminous United States, Environmental Management, 28(3), 293–316, doi:10.1007/s0026702429, 2001.

McManamay, R. A., Bevelhimer, M. S. and Kao, S.-C.: Updating the US hydrologic classification: an approach to clustering and stratifying ecohydrologic data: US HYDROLOGIC CLASSIFICATION, Ecohydrology, 7(3), 903–926, doi:10.1002/eco.1410, 2014.

Newman, A., Sampson, K., Clark, M., Bock, A., Viger, R. and Blodgett, D.: A large-sample watershed-scale hydrometeorological dataset for the contiguous USA, , doi:10.5065/D6MW2F4D, 2014.

Omernik, J. M. and Griffith, G. E.: Ecoregions of the Conterminous United States: Evolution of a Hierarchical Spatial Framework, Environmental Management, 54(6), 1249–1266, doi:10.1007/s00267-014-0364-1, 2014.

Pfister, L., Martínez-Carreras, N., Hissler, C., Klaus, J., Carrer, G. E., Stewart, M. K. and McDonnell, J. J.: Bedrock geology controls on catchment storage, mixing, and release: A comparative analysis of 16 nested catchments, Hydrological Processes, 31(10), 1828–1845, doi:10.1002/hyp.11134, 2017.

Potter, N. J., Zhang, L., Milly, P. C. D., McMahon, T. A. and Jakeman, A. J.: Effects of rainfall seasonality and soil moisture capacity on mean annual water balance for Australian catchments: WATER BALANCE OF AUSTRALIAN CATCHMENTS, Water Resources Research, 41(6), doi:10.1029/2004WR003697, 2005.

Samani, Z.: Estimating solar radiation and evapotranspiration using minimum climatological data, Journal of Irrigation and Drainage Engineering, 126(4), doi:https://doi.org/10.1061/(ASCE)0733-9437(2000)126:4(265), 2000.

Sawicz, K., Wagener, T., Sivapalan, M., Troch, P. A. and Carrillo, G.: Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA, Hydrology and Earth System Sciences, 15(9), 2895–2911, doi:10.5194/hess-15-2895-2011, 2011.

Schaake, J., Cong, S. Z. and Duan, Q. Y.: The US MOPEX data set, IAHS-AISH, 307, 9–28, 2006.

Singh, R., Archfield, S. A. and Wagener, T.: Identifying dominant controls on hydrologic parameter transfer from gauged to ungauged catchments – A comparative hydrology approach, Journal of Hydrology, 517, 985–996, doi:10.1016/j.jhydrol.2014.06.030, 2014.

Sivapalan, M.: Prediction in ungauged basins: a grand challenge for theoretical hydrology, Hydrological Processes, 17(15), 3163–3170, doi:10.1002/hyp.5155, 2003.

Sivapalan, M.: Pattern, Process and Function: Elements of a Unified Theory of Hydrology at the Catchment Scale, in Encyclopedia of Hydrological Sciences, edited by M. G. Anderson and J. J. McDonnell, John Wiley & Sons, Ltd, Chichester, UK., 2005.

Toth, E.: Catchment classification based on characterisation of streamflow and precipitation time series, Hydrology and Earth System Sciences, 17(3), 1149–1159, doi:10.5194/hess-17-1149-2013, 2013.

Trancoso, R., Phinn, S., McVicar, T. R., Larsen, J. R. and McAlpine, C. A.: Regional variation in streamflow drivers across a continental climatic gradient, Ecohydrology, 10(3), e1816, doi:10.1002/eco.1816, 2017.

Wagener, T., Sivapalan, M., Troch, P. and Woods, R.: Catchment Classification and Hydrologic Similarity, Geography Compass, 1(4), 901–931, doi:10.1111/j.1749-8198.2007.00039.x, 2007.

Ward, J. H.: Hierarchical Grouping to Optimize an Objective Function, Journal of the American Statistical Association, 58(301), 236–244, doi:10.1080/01621459.1963.10500845, 1963.

Westerberg, I. K. and McMillan, H. K.: Uncertainty in hydrological signatures, Hydrology and Earth System Sciences, 19(9), 3951–3968, doi:10.5194/hess-19-3951-2015, 2015.

Yaeger, M., Coopersmith, E., Ye, S., Cheng, L., Viglione, A. and Sivapalan, M.: Exploring the physical controls of regional patterns of flow duration curves – Part 4: A synthesis of empirical analysis, process modeling and catchment classification, Hydrology and Earth System Sciences, 16(11), 4483–4498, doi:10.5194/hess-16-4483-2012, 2012.

Yeung, K. Y. and Ruzzo, W. L.: Principal component analysis for clustering gene expression data, Bioinformatics, 17(9), 763–774, doi:10.1093/bioinformatics/17.9.763, 2001.

~~Addor, N., Newman, A. J., Mizukami, N. and Clark, M. P.: The CAMELS data set: catchment attributes and meteorology for large-sample studies, Hydrology and Earth System Sciences, 21(10), 5293–5313, doi:10.5194/hess-21-5293-2017, 2017.~~

~~Addor, N., Nearing, G., Prieto, C., Newman, A. J., Le Vine, N. and Clark, M. P.: A ranking of hydrological signatures based on their predictability in space, Water Resources Research, doi:10.1029/2018WR022606, 2018.~~

~~Ali, G., Tetzlaff, D., Soulsby, C., McDonnell, J. J. and Capell, R.: A comparison of similarity indices for catchment classification using a cross-regional dataset, Advances in Water Resources, 40, 11–22, doi:10.1016/j.advwatres.2012.01.008, 2012.~~

~~Alvarez-Garreton, C., Mendoza, P. A., Boisier, J. P., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., Lara, A., Puelma, C., Cortes, G., Garreaud, R., McPhee, J. and Ayala, A.: The CAMELS-CL dataset: catchment attributes and meteorology for large sample studies – Chile dataset, Hydrology and Earth System Sciences, 22(11), 5817–5846, doi:10.5194/hess-22-5817-2018, 2018.~~

~~Andréassian, V., Lerat, J., Le Moine, N. and Perrin, C.: Neighbors: Nature's own hydrological models, Journal of Hydrology, 414–415, 49–58, doi:10.1016/j.jhydrol.2011.10.007, 2012.~~

~~Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A. and Wood, E. F.: Present and future Köppen-Geiger climate classification maps at 1-km resolution, Scientific Data, 5, 180214, doi:10.1038/sdata.2018.214, 2018.~~

~~Berghuijs, W. R., Sivapalan, M., Woods, R. A. and Savenije, H. H. G.: Patterns of similarity of seasonal water balances: A window into streamflow variability over a range of time scales, Water Resources Research, 50(7), 5638–5661, doi:10.1002/2014WR015692, 2014.~~

~~Beven, K. J.: Uniqueness of place and process representations in hydrological modeling, Hydrology and Earth System Sciences, 4(2), 203–213, 2000.~~

~~Burn, D. H.: Catchment similarity for regional flood frequency analysis using seasonality measures, Journal of Hydrology, 202(1–4), 212–230, doi:10.1016/S0022-1694(97)00068-1, 1997.~~

~~Coopersmith, E., Yaeger, M. A., Ye, S., Cheng, L. and Sivapalan, M.: Exploring the physical controls of regional patterns of flow duration curves – Part 3: A catchment classification system based on regime curve indicators, Hydrology and Earth System Sciences, 16(11), 4467–4482, doi:10.5194/hess-16-4467-2012, 2012.~~

~~Costanza, R., Wainger, L., Folke, C. and Mäler, K.-G.: Modeling Complex Ecological Economic Systems: Toward an Evolutionary, Dynamic Understanding of People and Nature, in Ecosystem Management, pp. 148–163, Springer New York, New York, NY., 1993.~~

~~Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M. and Andréassian, V.: Large-sample hydrology: a need to balance depth with breadth, Hydrology and Earth System Sciences, 18(2), 463–477, doi:10.5194/hess-18-463-2014, 2014.~~

~~Haines, A. T., Finlayson, B. L. and McMahon, T. A.: A global classification of river regimes, Applied Geography, 8(4), 255–272, doi:10.1016/0143-6228(88)90035-5, 1988.~~

~~Jehn, F. U.: zutn/Catchment-Classification: Release for Zenodo, , doi:10.5281/zenodo.2203639, 2018.~~

~~Knoben, W. J. M., Woods, R. A. and Freer, J. E.: A Quantitative Hydrological Climate Classification Evaluated With Independent Streamflow Data, Water Resources Research, 54(7), 5088–5109, doi:10.1029/2018WR022913, 2018.~~

~~Kuentz, A., Arheimer, B., Hundecha, Y. and Wagener, T.: Understanding hydrologic variability across Europe through catchment classification, Hydrology and Earth System Sciences, 21(6), 2863–2879, doi:10.5194/hess-21-2863-2017, 2017.~~

~~Li, C., Cao, J., Nie, S.-P., Zhu, K.-X., Xiong, T. and Xie, M.-Y.: Serum metabolomics analysis for biomarker of Lactobacillus plantarum NCU116 on hyperlipidaemic rat model feed by high fat diet, Journal of Functional Foods, 42, 171–176, doi:10.1016/j.jff.2017.12.036, 2018.~~

~~McGlynn, B. L. and Seibert, J.: Distributed assessment of contributing area and riparian buffering along stream networks: TECHNICAL NOTE, Water Resources Research, 39(4), doi:10.1029/2002WR001521, 2003.~~

~~McMahon, G., Gregonis, S. M., Waltman, S. W., Omernik, J. M., Thorson, T. D., Freeouf, J. A., Rorick, A. H. and Keys, J. E.: Developing a Spatial Framework of Common Ecological Regions for the Conterminous United States, Environmental Management, 28(3), 293–316, doi:10.1007/s0026702429, 2001.~~

~~McManamay, R. A., Bevelhimer, M. S. and Kao, S.-C.: Updating the US hydrologic classification: an approach to clustering and stratifying ecohydrologic data: US HYDROLOGIC CLASSIFICATION, Ecohydrology, 7(3), 903–926, doi:10.1002/eco.1410, 2014.~~

~~Newman, A., Sampson, K., Clark, M., Bock, A., Viger, R. and Blodgett, D.: A large-sample watershed-scale hydrometeorological dataset for the contiguous USA, , doi:10.5065/D6MW2F4D, 2014.~~

~~Potter, N. J., Zhang, L., Milly, P. C. D., McMahon, T. A. and Jakeman, A. J.: Effects of rainfall seasonality and soil moisture capacity on mean annual water balance for Australian catchments: WATER BALANCE OF AUSTRALIAN CATCHMENTS, Water Resources Research, 41(6), doi:10.1029/2004WR003697, 2005.~~

~~Sawicz, K., Wagener, T., Sivapalan, M., Troch, P. A. and Carrillo, G.: Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA, Hydrology and Earth System Sciences, 15(9), 2895–2911, doi:10.5194/hess-15-2895-2011, 2011.~~

~~Schaake, J., Cong, S. Z. and Duan, Q. Y.: The US MOPEX data set, IAHS-AISH, 307, 9–28, 2006.~~

~~Semenova, O. and Beven, K.: Barriers to progress in distributed hydrological modelling: Invited Commentary, Hydrological Processes, 29(8), 2074–2078, doi:10.1002/hyp.10434, 2015.~~

~~Singh, R., Archfield, S. A. and Wagener, T.: Identifying dominant controls on hydrologic parameter transfer from gauged to ungauged catchments – A comparative hydrology approach, Journal of Hydrology, 517, 985–996, doi:10.1016/j.jhydrol.2014.06.030, 2014.~~

~~Sivapalan, M.: Prediction in ungauged basins: a grand challenge for theoretical hydrology, Hydrological Processes, 17(15), 3163–3170, doi:10.1002/hyp.5155, 2003.~~

~~Sivapalan, M.: Pattern, Process and Function: Elements of a Unified Theory of Hydrology at the Catchment Scale, in Encyclopedia of Hydrological Sciences, edited by M. G. Anderson and J. J. McDonnell, John Wiley & Sons, Ltd, Chichester, UK., 2005.~~

~~Toth, E.: Catchment classification based on characterisation of streamflow and precipitation time series, Hydrology and Earth System Sciences, 17(3), 1149–1159, doi:10.5194/hess-17-1149-2013, 2013.~~

~~Trancoso, R., Phinn, S., McVicar, T. R., Larsen, J. R. and McAlpine, C. A.: Regional variation in streamflow drivers across a continental climatic gradient, Ecohydrology, 10(3), e1816, doi:10.1002/eco.1816, 2017.~~

~~Troch, P. A., Carrillo, G., Sivapalan, M., Wagener, T. and Sawicz, K.: Climate-vegetation-soil interactions and long-term hydrologic partitioning: signatures of catchment co-evolution, Hydrology and Earth System Sciences, 17(6), 2209–2217, doi:10.5194/hess-17-2209-2013, 2013.~~

~~Wagener, T., Sivapalan, M., Troch, P. and Woods, R.: Catchment Classification and Hydrologic Similarity, Geography Compass, 1(4), 901–931, doi:10.1111/j.1749-8198.2007.00039.x, 2007.~~

~~Ward, J. H.: Hierarchical Grouping to Optimize an Objective Function, Journal of the American Statistical Association, 58(301), 236–244, doi:10.1080/01621459.1963.10500845, 1963.~~

~~Westerberg, I. K. and McMillan, H. K.: Uncertainty in hydrological signatures, Hydrology and Earth System Sciences, 19(9), 3951–3968, doi:10.5194/hess-19-3951-2015, 2015.~~

~~Yaeger, M., Coopersmith, E., Ye, S., Cheng, L., Viglione, A. and Sivapalan, M.: Exploring the physical controls of regional patterns of flow duration curves – Part 4: A synthesis of empirical analysis, process modeling and catchment classification, Hydrology and Earth System Sciences, 16(11), 4483–4498, doi:10.5194/hess-16-4483-2012, 2012.~~

~~Yeung, K. Y. and Ruzzo, W. L.: Principal component analysis for clustering gene expression data, Bioinformatics, 17(9), 763–774, doi:10.1093/bioinformatics/17.9.763, 2001.~~

**Appendix**

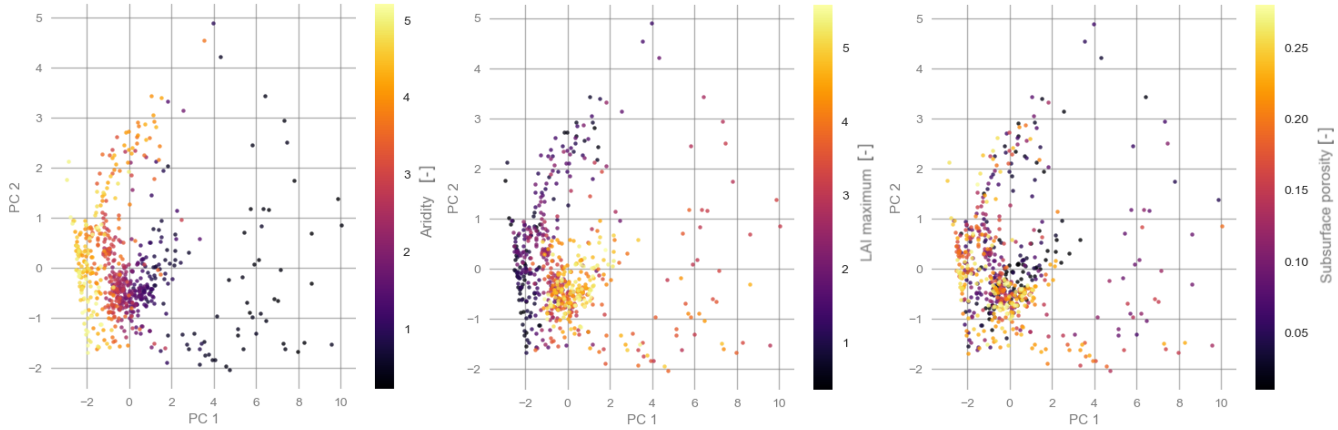



**Figure A1: Patterns of catchment attributes in the PCA space of the hydrological signatures, with decreasing strength of the observed pattern from left (aridity) to right (subsurface porosity).~~.~~**

~~**A 1.1 Detailed description of the catchment clusters**~~

~~**Cluster 1** is defined by a high cover of vegetation. In addition, most catchments are located at low elevations, experience little snow and have a deep bedrock. Hydrologically these catchments have little discharge. They are mainly located in the Southeastern and Central Plains and therefore get relative high rainfall (> 1000 mm year). Their low discharge is probably caused by the low elevation those catchments are located, groundwater discharge and the high evaporation of the forests. Cluster 1 also contains the largest amount of catchments from all cluster (n = 230). So over one third of the catchments in CAMELS show a relatively similar behavior.~~

~~**Cluster 2** most typical attribute in comparison with the other catchments is its depth to the bedrock. However, concerning the catchment attributes cluster 2 is undefined as it contains catchments of most regions of the continental United States (with a focus on the Central Plains). The hydrological signatures on the other hand show a clearer pattern. Here, the mean winter discharge, Q95 and the mean annual discharge have a narrow range. This shows that catchments with very different attributes can produce very similar discharge characteristics, as the different attributes seems to cancel each other out in their influence on the discharge.~~

~~**Cluster 3** is the smallest cluster with only seven catchments. Those are all located in the Northwestern Forested Mountains. Their most distinct feature is their uniform high cover with forest. They also experience high precipitation events only seldom and precipitation is snow half of the time. Hydrologically their most distinct features is their very high mean summer discharge and high runoff ratio, which is probably caused by the large amounts of snow these catchments receive.~~

~~**Cluster 4** is also located in the Northwestern Forested Mountains, with the exception of four catchments that are located in Florida. This again is an example of different catchment attributes being able to create similar discharge characteristics concerning their signatures, while having different catchment attributes. The catchments have overall low discharge and few high flow events, while their catchment attributes vary widely, especially in all attributes that are related to elevation (e.g. fraction of precipitation falling as snow).~~

~~**Cluster 5,** has only few catchments (n = 9). They are all located at regions in the northern part of the Marin West Coast Forests. This is the region in the continental US that receives the highest precipitation (> 2000 mm year). This is mirrored in their discharge characteristics. These catchments have the highest discharge in the whole dataset, especially in the summer. They are also uniformly covered by almost 100 % of forest. They also experience only few high precipitation events as they get rain and snow more or less constantly in the same amount.~~

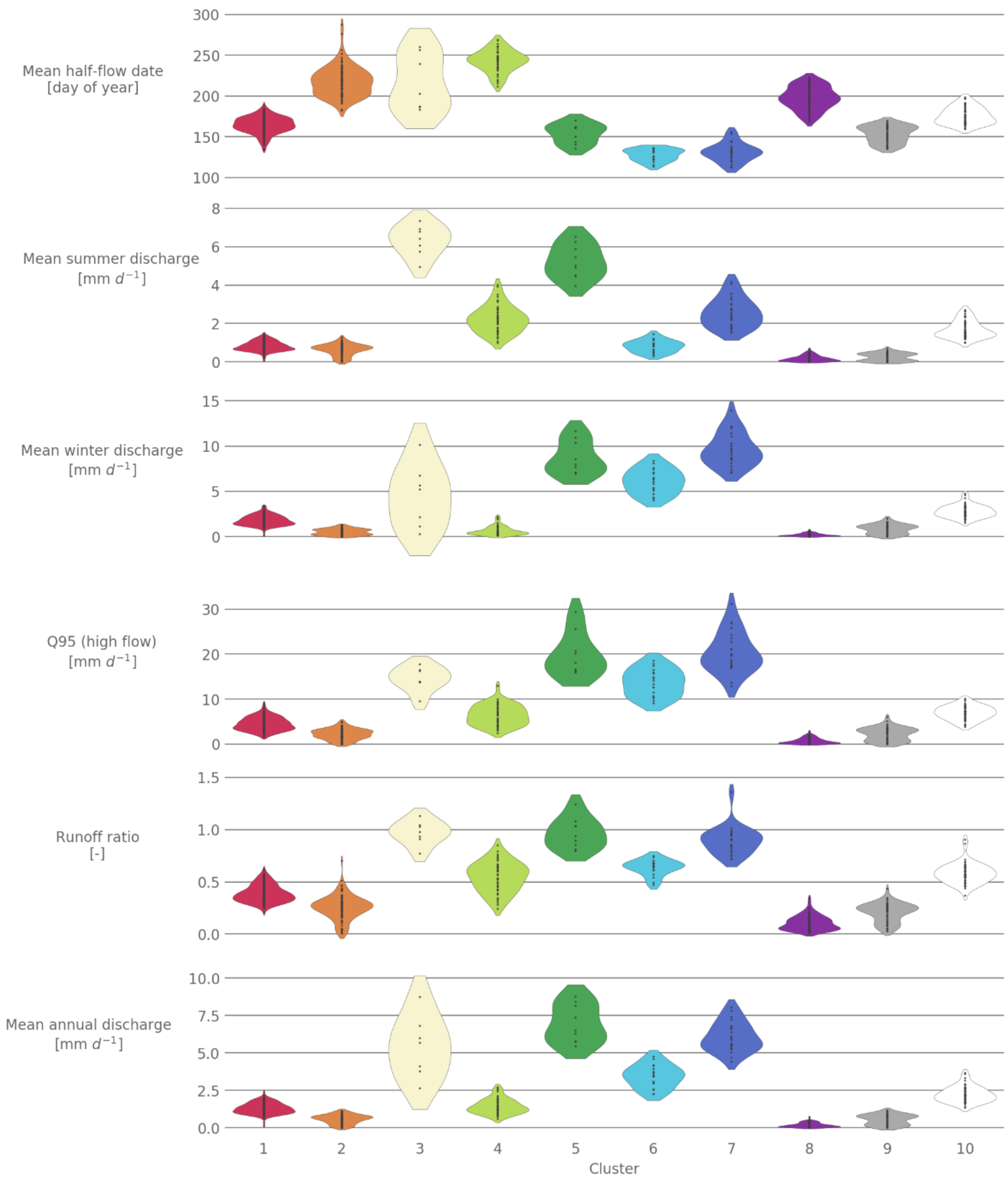
~~**Cluster 6** catchments are also located in the Marine West Coast Forest, but cover the whole region and not only the northern part like Cluster 5. The catchments are very similar in their attributes and discharge characteristics to Cluster 5, with the exception of a lower discharge and runoff ratio. This might be caused by a slightly lower precipitation in comparison with Cluster 5.~~

~~**Cluster 7** is also located in the same region as Cluster 5 and 6 (Marine West Coast Forests). Concerning the catchment attributes and the discharge characteristics, it is located between Cluster 5 and 6. So, Cluster 5 to 7 all cover the same region and differ in their mean summer discharge, which is caused by slight variations in elevation and location.~~

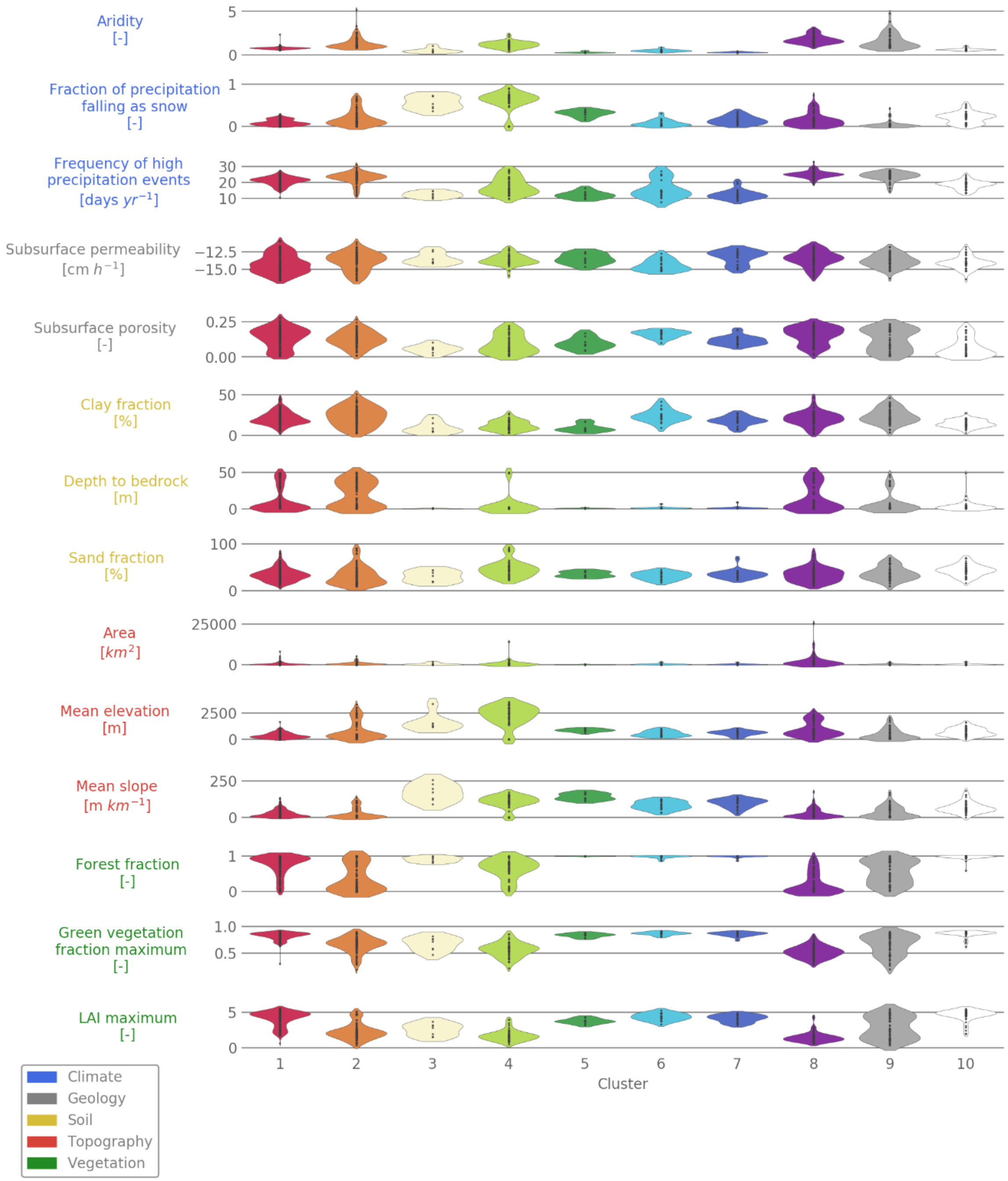
~~**Cluster 8** is the overall most arid cluster catchments. All of the catchments are located in western parts of the Great Plains and in the North American Deserts. They are shaped by an overall little availability of water and high evaporation, which is shown in the very low mean annual discharge and runoff ratio. This also results in low values for the LAI. However, the frequency of high precipitation events is high.~~

~~**Cluster 9** covers all southern states of the United States. The catchments here are quite similar to Cluster 8, but show a lower seasonality (as indicated by their lower half flow date) and a higher forest cover and green vegetation.~~

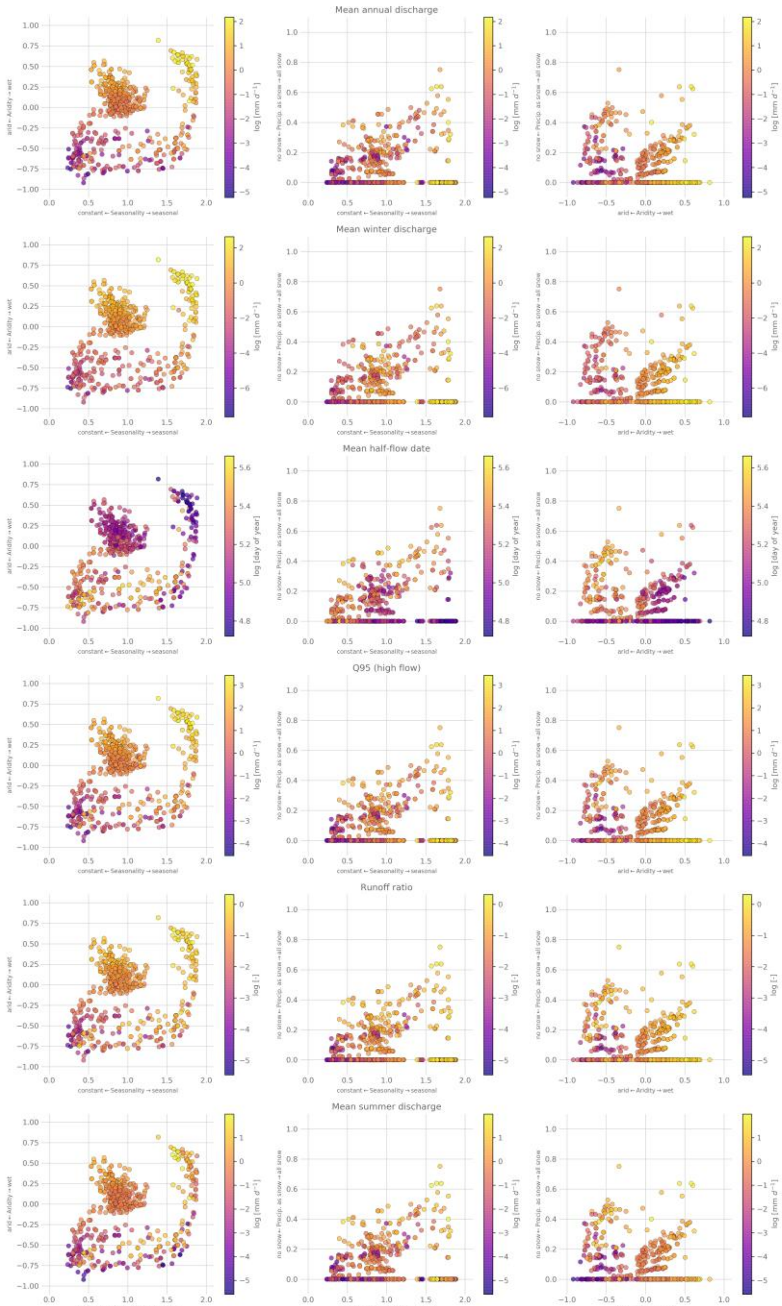
~~**Cluster 10** catchments are located in the Appalachian Mountains. The mean elevation higher than most other clusters and the catchments also have low aridity and a very high forest cover. Their discharge characteristics is similar to the Marine West Coast Forests of Cluster 5 to 7. However, they receive less water than those catchments and experience a higher seasonality (as indicated by the higher mean half-flow date).~~




~~**Figure A2: Violin plot of the hydrological signatures sorted by catchment clusters. Single dots in the violins indicate the single catchments.**~~




~~Figure A3: Violin plots of the catchment attributes sorted by catchment clusters. Single dots in the violins indicate the single catchments.~~




895 **Figure A42: Hydrological signatures for all catchments in the climate index space of Knoben et al. (2018). Single dots show the catchments and are colored according to the value of the mean annual discharge. The log of the signatures is used to show the relative differences between the catchments.**