Major changes in the manuscript:

1) Title: As major parts of the paper were rewritten and we did additional analyses, the old title did not capture the topic of the paper anymore. Therefore, we changed the title to reflect those changes (in agreement with the editor).
2) Abstract: To accommodate the major changes, we have completely rewritten the abstract.
3) Results and Discussion:
    a. Changed order of subsections, to allow easier reading.
    b. Complete rewrite of subsection 3.3 (Exploration of the catchment clusters). We now discuss the clusters in more detail and provide additional analyses concerning the location of the catchments and their catchment clusters. In addition, we also discuss differences in catchment attributes in the East and West of the US.
    c. Subsection 3.4 (Differences in clusters in comparison with other hydrological clustering studies) discusses other clustering studies in more detail.
    d. Added the new section 3.5 (Comparing catchment clusters based on hydrological behavior and climate). In this subsection, we compare our clustering using hydrological signatures with climatic clusters and discuss the findings.
4) Summary and conclusion: To accommodate the major changes, we have completely rewritten the summary and conclusion.
5) Appendix: Added Figure A4, which depicts all hydrological signatures used in the climate index space of Knoben et al. (2018).

We would like to thank the reviewers for their constructive comments on the manuscript "Clustering CAMELS using hydrological signatures with high spatial predictability"

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

Response to Reviewer #1 (Anonymous)

Jehn et al. classified the CAMELS catchments based on hydrological signatures, and subsequently investigated the link between catchment attributes and the classes. The conclusion of the study is that catchment behavior can mainly be attributed to climate in regions with homogeneous topography, but that this is more difficult in regions with heterogeneous topography. Unfortunately, my perception is that the conclusions of the study are based on a fallacy. The main problem can be found here: "If climate were the main driver, the clusters would be located along a climatic gradient. However, this is only true for the eastern half of the United States (for a climatic map of the United states see (Beck et al.,2018). In this part of the United States, the low relief allows large regions with a uniform climate, that only changes of larger scales." If looking at the map in Beck et al. (2018), but also Peel et al. (HESS, 2007), or Knoben et al. (WRR, 2018), indeed the eastern part of the US shows large regions with uniform climate. But the maps also all show the large scattering in climates in the west: there is more spatial variation in climate in the western part of the US. This therefore seems no justification to state that climate is less relevant in regions with varying topography - there, the climate is just more variable too. This is also confirmed by the results of the study, where precipitation falling as snow is found as one of the main indicators in the west. "This implies that climate is a good indicator for the discharge characteristics as long as the topography is homogenous." seems therefore a too strict and incorrect conclusion, that does not necessarily follows from the results / figures.

After receiving those very constructive comments on our first version of the manuscript, we did a mayor reanalysis of our data and provide additional comparisons and tests. This changed our conclusions and we now have rewritten substantial parts of the paper to accommodate this. Section 3.3 was in the focus of the reviewers concerns and is now changed to:

3.3 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). 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.

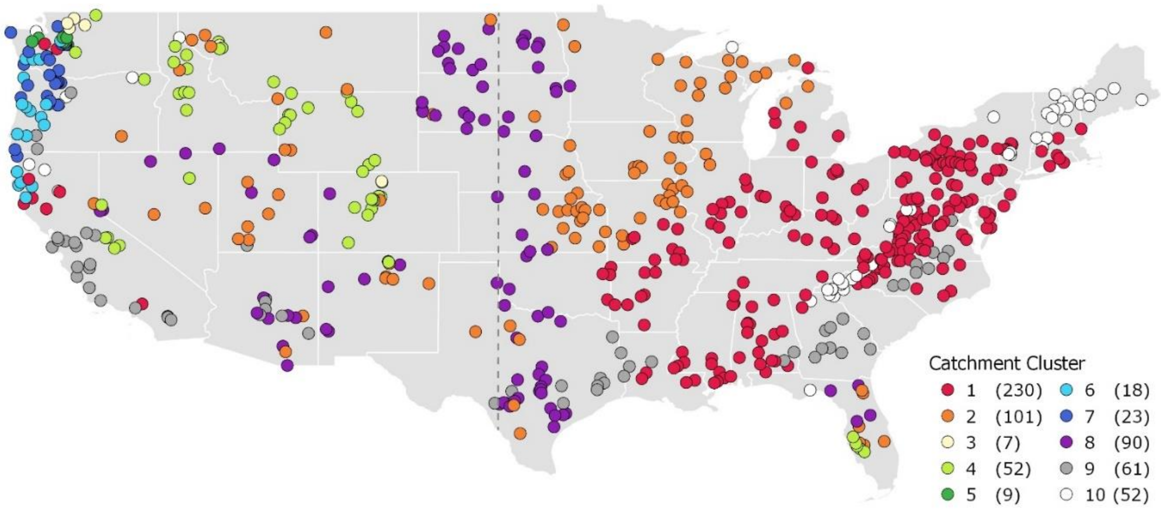


Figure 3: Locations of the clustered CAMELS catchments in the continental US. Dotted line marks the 100th meridian.

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 West, while Cluster 1 and 10 are in the East. However, the remaining Clusters 2, 8 and 9 have roughly similar amounts of catchments in both regions. 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.

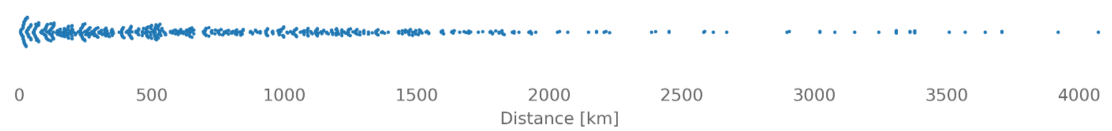


Figure 4: Swarm plot of the real world distances of all catchments to the most hydrologically similar catchment (based on their distance in the PCA space of the hydrological signatures).

In addition, similar catchments can be quite far away from each other (Figure 4). Sometimes, the catchment with the most similar signature was found as far as 4000 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 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 catchment 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 obviously different climates, can lead to similar (equifinal) discharge behavior.

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.

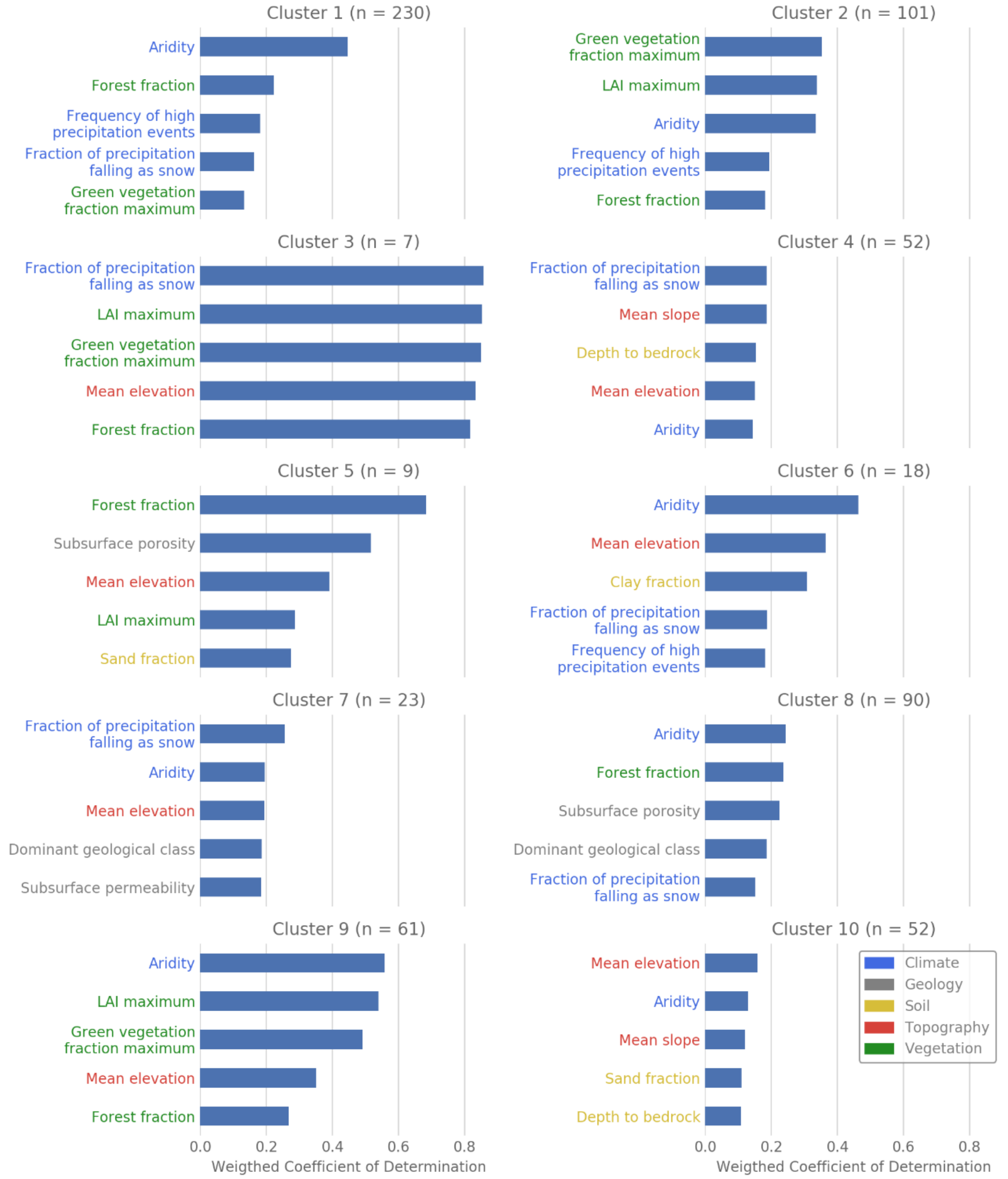


Figure 5: Importance of the catchment attributes evaluated by the quadratic regression. For the catchment clusters. Attributes colored according to their catchment attribute class.

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, 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)*

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 R^2 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 R^2 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.

Overall, 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 a less strong climatic signal in those regions. 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- and catchment-attributes data cannot sufficiently explain the hydrological characteristics.*

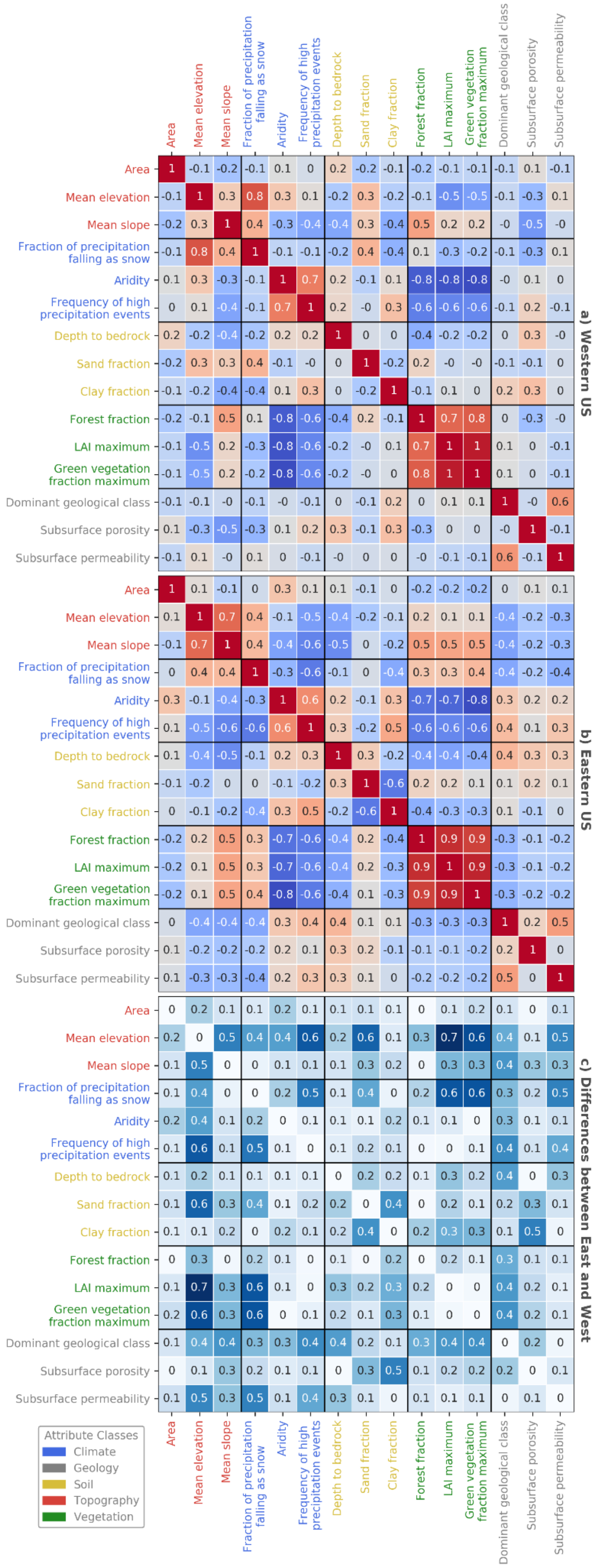


Figure 6: Correlation of all catchment attributes for western (a) and eastern (b) US and absolute differences in correlation between the eastern and western US. Eastern and western is defined by the 100th meridian.

Furthermore, I wonder to what extent 'homogeneous topography' can be found as criterion, when looking at the catchment scale, considering that most catchments in CAMELS are rather small.

We do not use this phrasing anymore in the revised version of the manuscript.

Besides my disagreement with the main conclusion, I consider the insights gained from the paper low compared to already available literature, especially considering Addor etal. (2018) and Knoben et al. (WRR, 2018). What did we learn from this study about the relation between attributes and signatures, or catchment clustering, that was unknown before? Especially given that I disagree with the main conclusion. If is it the method applied (PCA combined with clustering), then further elaborate on the methods and better explain everything that is done and how it differs from other studies. This also needs explanation why this method would provide insights that cannot / haven't been obtained with other methods. I would like to encourage the authors to dive deeper into the material and expand the analysis, and have a critical look at their own conclusions.

To widen the scope of this study and to address the differences in clustering approaches that use hydrological behavior and climate respectively, we added a new section to discuss those topics:

3.5 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) 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 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)).

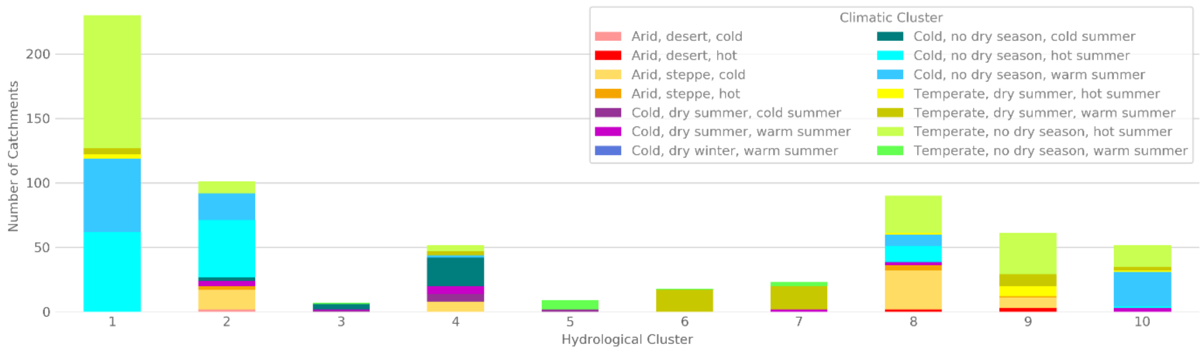


Figure 7: Membership of Koeppen-Geiger clusters (Beck et al. (2018)) in the hydrological clusters.

The picture is less clear concerning the climatic index space of Knoben et al. (2018) (Figure 8a). 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. 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, we could not identify a clear dominating catchment attribute, if we look at Cluster 4 (located in the Northwestern Forested Mountains and Florida) (Figure 5). Catchments with this hydrological behavior can be found in the space of the climatic indices of Knoben et al. 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 8b, A4) 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. This is even more remarkable, as the signatures used are linked to climate directly. For example, the signature "mean half flow date" can be seen as a measure of seasonality. 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.

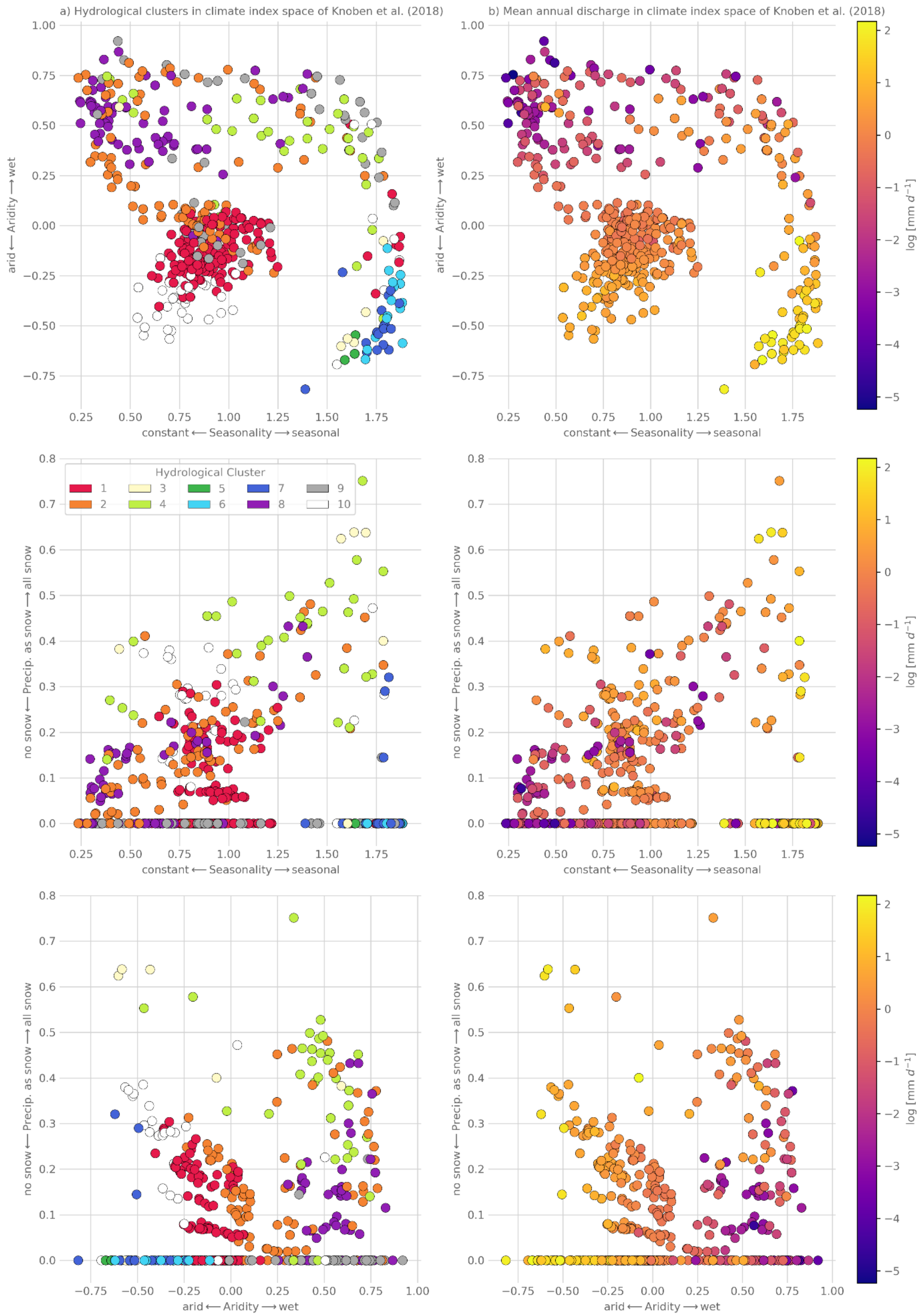


Figure 8: a) Comparison of the hydrological clustering of this study with the climate index space of Knoben et al. (2018). Single dots show the catchments and are colored by their hydrological clusters. b) Mean annual discharge for all catchments in the climate index space of Knoben et al. (2018). Single 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 A4.

To reflect the abovementioned changes, we have also rewritten the abstract and the summary and conclusion:

- 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 analyze the connections between the resulting clusters and the catchment attributes and relate this to the co-variability of the catchment attributes. 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 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 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 conclusion

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 catchment 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, 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.

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. Nevertheless, 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, it might be a worthwhile pathway to include measures of spatial variability of the climate in the single catchments. 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.*

We would like to thank the reviewers for their constructive comments on the manuscript "Clustering CAMELS using hydrological signatures with high spatial predictability"

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

Response to Reviewer #2 (Andrew Newman)

General comments: This paper examines the CAMELS catchments and clusters them using hydrologic signatures that have been previously found to have high spatial predictability. Overall this study is somewhat unsatisfying. Little new physical insight is gained in understanding how we determine similarity across catchments. The results do agree with past studies, which is a good test of the previous work. However, what does this specific study bring to us? Previous results discussed here, e.g. Addor et al. (2018) (Fig. 4), and Newman et al. (2015) (Fig. 12) found the same results. Aridity is the primary driver of basin behavior given the catchment scale attributes used, followed by other climate indices (e.g. snow). Finally, one of the primary conclusions drawn from the clustering results needs to be reexamined (specific comment #2).

First of all, we would like to thank the Reviewer herein to provide his very constructive comments. We tried to pick up all points, which lead to a revised version of this manuscript, which provides from our point of view now a clearer insight into the gained understanding and the novelty of this research.

Specific comments:

1) Is 10 clusters necessary? Why does 10 make this study similar to others? Did those studies arbitrarily pick 10 also? I wonder if a similar cluster selection method is better, rather than the same number of clusters. A more detailed justification in the methods section is necessary.

To further elaborate on our choice of 10 clusters, we added a more detailed explanation of our decision in section 2.3:
From those studies, Kuentz et al. (2018) 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. (2018), we also used ten clusters.

2) Many of the attributes have high co-variability. For example, elevation and temperature/fraction of snowfall, elevation and mean slope, forest fraction and elevation (in the western US) are likely candidates. Addor et al. (2018) discusses this briefly, but much more could be done here. It would be good to understand this co-variability and modify the discussion accordingly, particularly the conclusions on lines 173-180. Spatial proximity or the attributes defined as climate by the authors are bad predictors in areas with heterogeneous topography precisely because topography and climate are intertwined. That does not mean that climate is a poor predictor of catchment behavior in those same regions.

In light of the reanalysis of our data, we have mostly rewritten section 3.3 and added a discussion of the co-variability of the catchment attributes. This also changes our discussion of the connection between the topography and the climate.

3.3 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). 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.

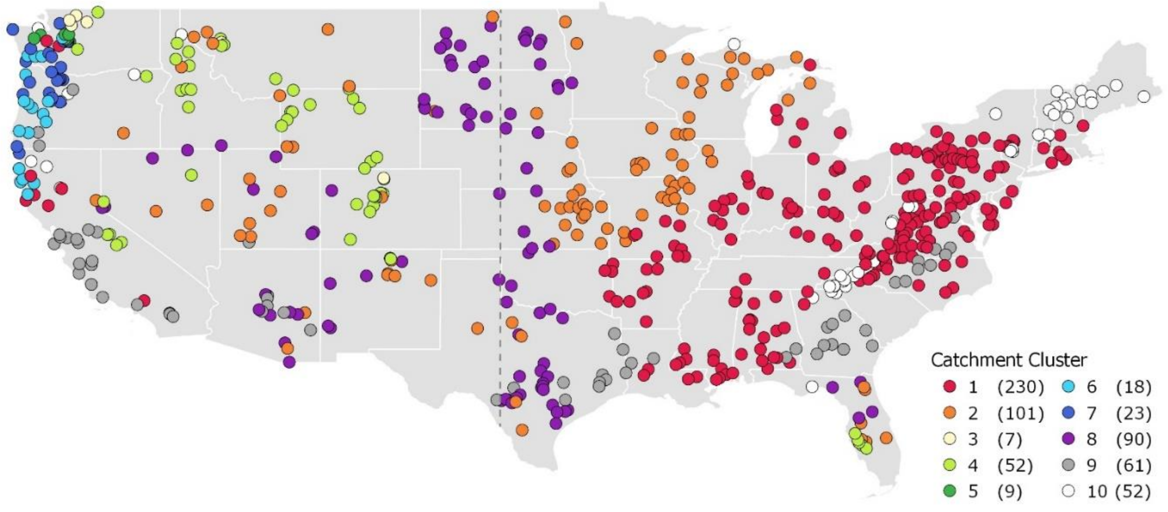


Figure 3: Locations of the clustered CAMELS catchments in the continental US. Dotted line marks the 100th meridian.

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 West, while Cluster 1 and 10 are in the East. However, the remaining Clusters 2, 8 and 9 have roughly similar amounts of catchments in both regions. 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 4: 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 4). Sometimes, the catchment with the most similar signature was found as far as 4000 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 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 catchment 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 obviously different climates, can lead to similar (equifinal) discharge behavior.*

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.

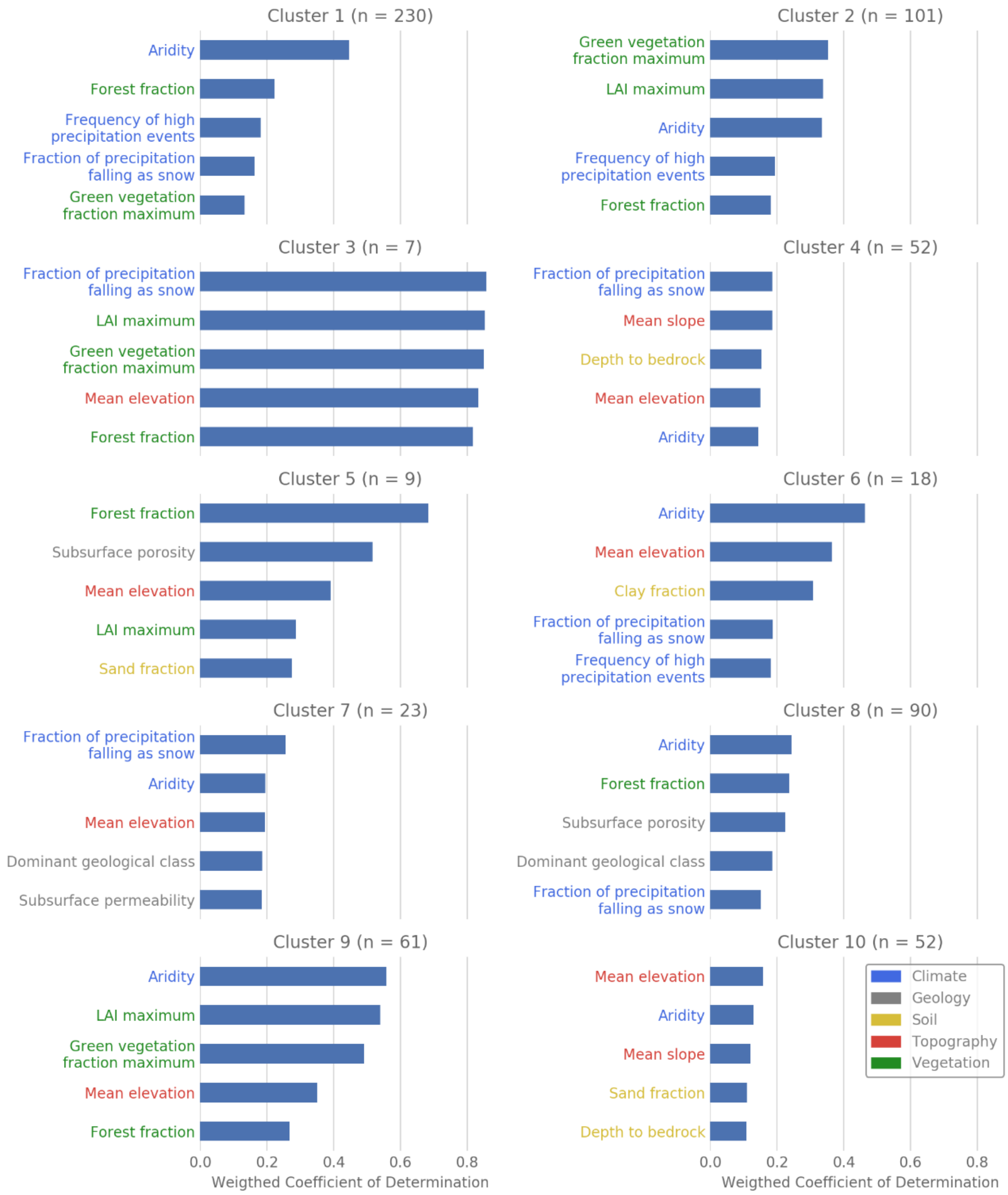


Figure 5: Importance of the catchment attributes evaluated by the quadratic regression. For the catchment clusters. Attributes colored according to their catchment attribute class.

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, 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)*

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 R^2 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 R^2 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.

Overall, 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 a less strong climatic signal in those regions. 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- and catchment-attributes data cannot sufficiently explain the hydrological characteristics.*

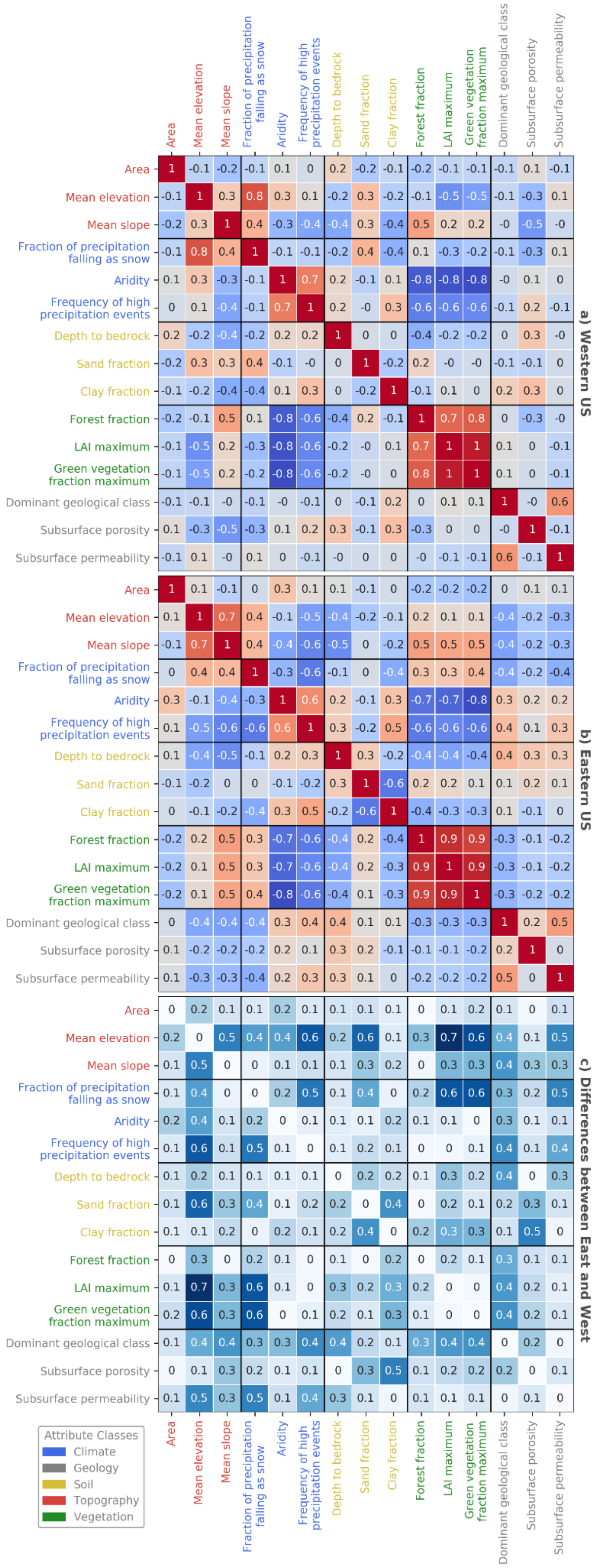


Figure 6: Correlation of all catchment attributes for western (a) and eastern (b) US and absolute differences in correlation between the eastern and western US. Eastern and western is defined by the 100th meridian.

3) Could more explanation be given as to how the clusters contain basins from very different locations (e.g. cluster 4)? There is some discussion in the appendix, which is good, but this cluster highlights limitations in our current clustering methods or application of those methods. How could other hydrologic signatures be used to provide more discriminatory power? Is predictability in space the best metric to determine which signatures to use in a study like this?

This is now discussed in section 3.3:

[…]

This indicates that clustering by using spatial proximity might only work in regions like in the eastern US, where the behavior of rivers 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 catchment quite far away from each other (e.g. Cluster 4). The catchments might be far away from each other, but the interplay of different catchment attributes and driving factors can lead to similar discharge behavior

[…]

Also, it seems like more discussion on the issues/benefits of using this method (clustering on principle components) using already aggregated data (signatures and catchment averaged attributes) would be useful. This could help the community learn more from these various clustering studies. The authors already provide a summary discussion relating these results to other studies, so I do not feel like this is out of scope.

We added a short discussion of this to section 3.4:
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). Those clusters also show distinct spatial patterns (Figure 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.

Minor comments:

The sentence starting on line 55 and ending on line 59 is a very long run-on sentence. It is hard to follow and should be reworked. I suggest checking the manuscript for other instances of run-on sentences.

Changed as proposed.

Figures: 1) Please consider increasing the contrast in the cluster colors in Figures 1and 3. Specifically clusters 1-3, and 4-6 are hard to visually separate.

Changed as proposed. We changed all figures with clusters to a more easily distinguishable color scheme. We also changed Figure A2 and A3 from swarm plots to violin plots, to make them easier to interpret.

~~Clustering CAMELS using hydrological signatures with high spatial predictability~~

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 analyze the connections between the resulting clusters and the catchment attributes and relate this to the co-variability of the catchment attributes. 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 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 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.~~The behavior of every catchment is unique. Still, we need ways to classify them as this helps to improve hydrological theories. Usually catchments are classified along either their attributes classes (e.g. climate, topography) or their discharge characteristics, which is often captured in hydrological signatures. However, recent studies have shown that many hydrological signatures have a low predictability in space and therefore only dubious hydrological meaning. Therefore, this study uses hydrological signatures with the highest predictability in space to cluster 643 catchments from the continental United States~~

~~(CAMELS (Catchment Attributes and MEteorology for Large-Sample Studies) dataset) into ten groups. We then evaluated the connection between catchment attributes with the hydrological signatures with quadratic regression, both in the overall CAMELS dataset and the ten clusters. In the overall dataset, aridity had the strongest connection to the hydrological signatures, especially in the eastern United States. However, the clusters in the western United States showed a more heterogeneous pattern with a larger influence of forest fraction, the mean elevation or the snow fraction. From this, we conclude that catchment behavior can be mainly attributed to climate in regions with homogenous topography. In regions with a heterogeneous topography, there is no clear pattern of the catchment behavior, as catchments show high spatial variability in their attributes. The classification of the CAMELS dataset with the hydrological signatures allows testing hydrological models in contrasting environments.~~

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 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.

Therefore, we selected the best six hydrological signatures with spatial predictability to classify catchments of the CAMELS (Catchment Attributes and M~~e~~Eteorology for Large-Sample Studies) dataset (Addor et al., 2017). Those six hydrological signatures are evaluated together with the fifteen 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. ~~, which can be used for further research. In addition, it will address the question, if the hydrological behavior is influenced from different catchment attributes, on the scale of the individual clusters and the whole dataset, respectively.~~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 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 (Table 1). 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 used three attributes per class. *Climate*: aridity, frequency of high precipitation events, fraction of precipitation falling as snow; *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 ~~together with their description~~ (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}

2.3 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). We only used principal components that account for at least 80% of the total variance of the hydrological signatures ~~similar to Kuentz et al. (2017)~~, which resulted in two principal components. We evaluated the connection between the principal components and the catchments attributes with the following procedure:

1) ~~First~~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 ~~coefficient of determination~~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 principal components were subsequently added, to obtain one coefficient of regression for every catchment attribute.

Quadratic regression was selected as interactions in natural hydrological systems are known to have unclear patterns and cannot be fitted with a 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).

The principal components 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). From those studies, Kuentz et al. (2018) 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. (2018), we also used ten clusters. ~~To make our results comparable to other published studies like Kuentz et al. (2017), we split the dataset into ten clusters.~~

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.

3 Results and Discussion

~~3.1 Relation of the principal components and the hydrological signatures~~

~~The rivers considered in this study show a wide range in hydrological signatures. This can be seen in the clusters of principal components of the hydrological signatures (Figure 1). However, most of the rivers are opposite of the loading vectors (the loading vectors are shown as arrows in the figure). 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 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.~~

~~Figure 1: Biplot of the principal components (PC). Colors indicate the cluster of the catchment.~~

3.1~~2~~ Impacts of catchment attributes on discharge characteristics in the whole dataset

~~After the clustering, w~~First we examined the weighted ~~coefficient of determination~~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 ~~coefficient of determination~~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.

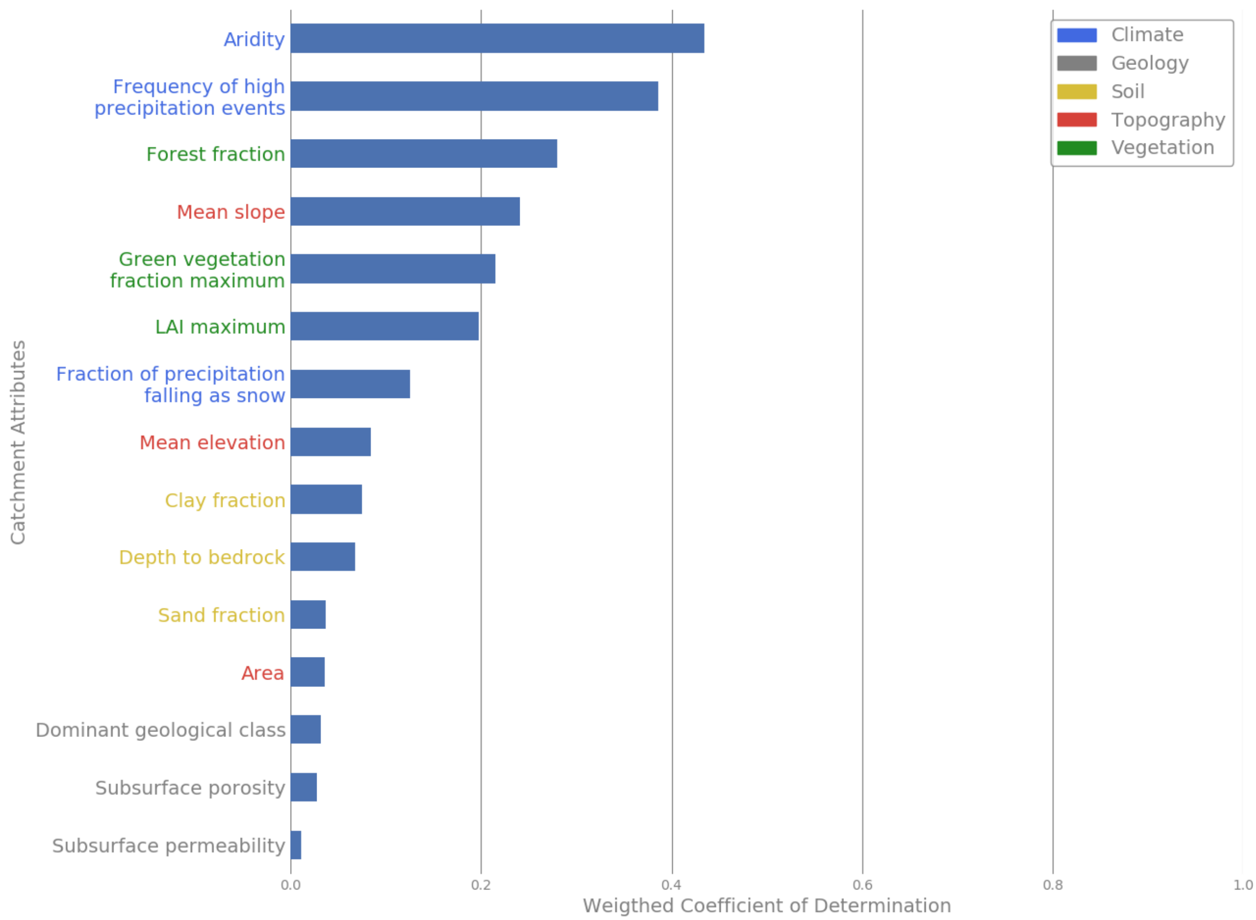


Figure 1: Importance of catchment attributes evaluated by quadratic regression for all considered catchments. Attributes colored according to their catchment attribute class.

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~~. (Table 1)~~. Nevertheless, our study probably captures a more general trend as we used a larger dataset and more hydrologically meaningful hydrological signatures ~~which have a better predictability 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. However, Coopersmith et al. (2012) only analyzed the flow duration curve, which has a mediocre predictability in space and it is therefore more unclear what it really depicts (Addor et al., 2018). Overall, this study here is in line with other literature in the field. Using the weighted ~~coefficient of determination~~R^2 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 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).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.

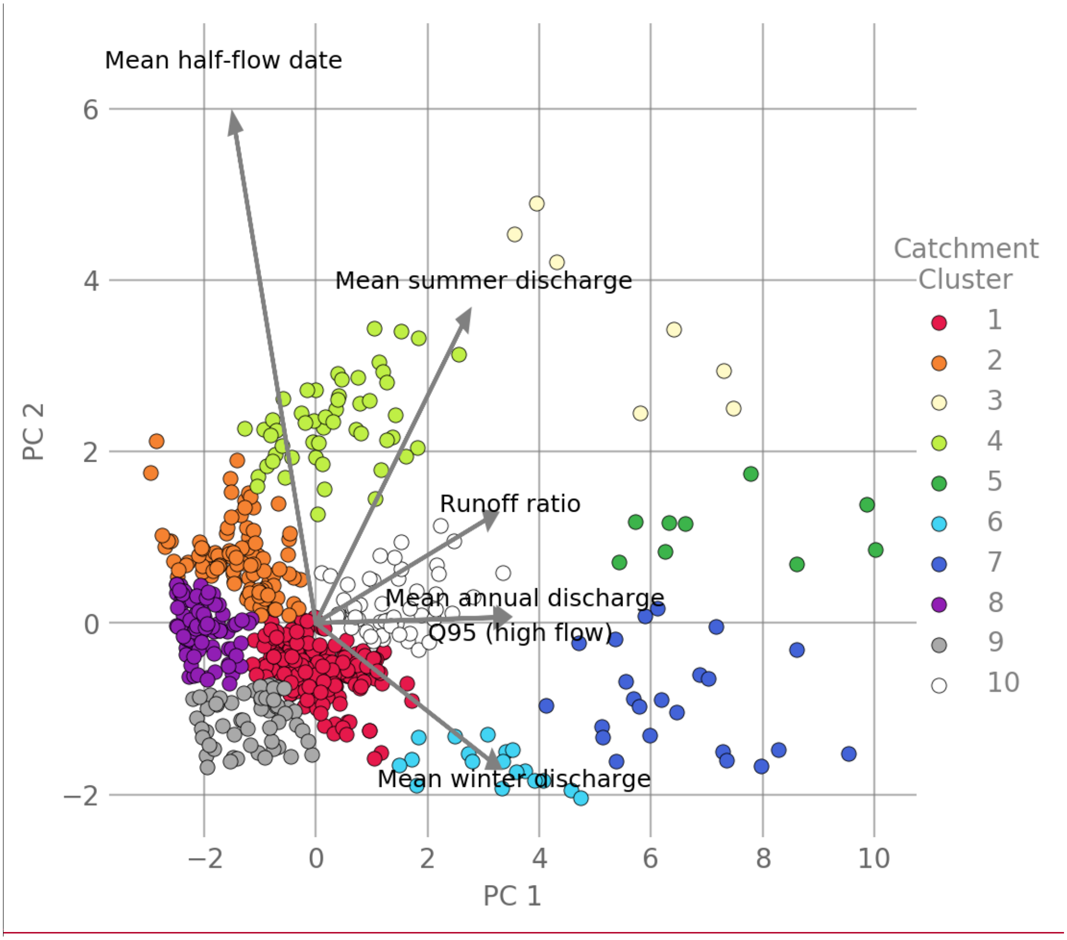


Figure 2: Biplot of the principal components (PC). Colors indicate the cluster of the catchment.

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

3.3 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). 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. ~~While the catchment attributes in the CAMELS and other datasets, as a whole, 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 catchment clusters. This is directly observable in the spatial distribution of the clusters (Figure 3). If climate were the main driver, the clusters would be located along a climatic gradient. However, this is only true for the eastern half of the United States (for a climatic map of the United states see (Beck et al., 2018). In this part of the United States, the low relief allows large regions with a uniform climate, that only changes of larger scales.~~

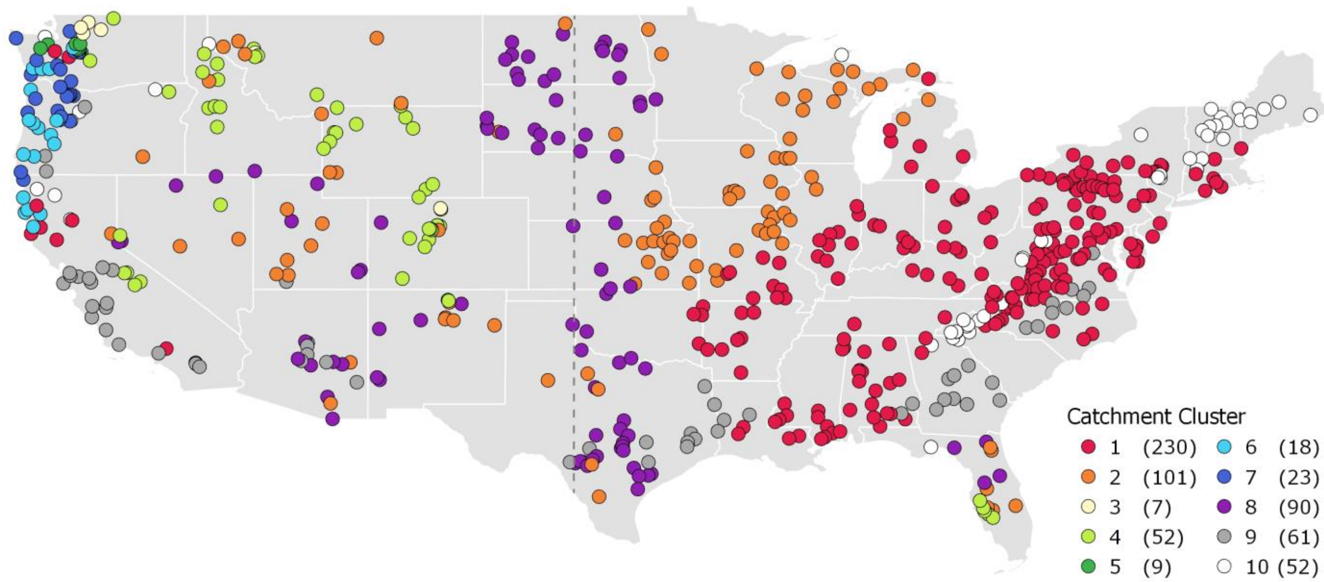


Figure 3: Locations of the clustered CAMELS catchments in the continental US. Dotted line marks the 100th meridian.

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 West, while Cluster 1 and 10 are in the East. However, the remaining Clusters 2, 8 and 9 have roughly similar amounts of catchments in both regions. 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 4: 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 4). Sometimes, the catchment with the most similar signature was found as far as 4000 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 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 catchment 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 obviously different climates, can lead to similar (equifinal) discharge behavior.

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.

~~The analysis of the importance of the catchment attributes in the clusters shows a different picture than for the whole dataset (compare Figure 2 and Figure 4). For Cluster 1 (Southeastern and Central Plains), 6 (Marine West Coast Forests), 8 (Great Plains and Deserts) and 9 (Southern states) aridity still has the clearest connection to the clusters. However, this is not the case for the remaining catchment clusters. Here the most important catchment attributes differ from cluster to cluster. 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 coefficient of determination, which is almost as high as the one for the fraction of precipitation falling as snow.~~

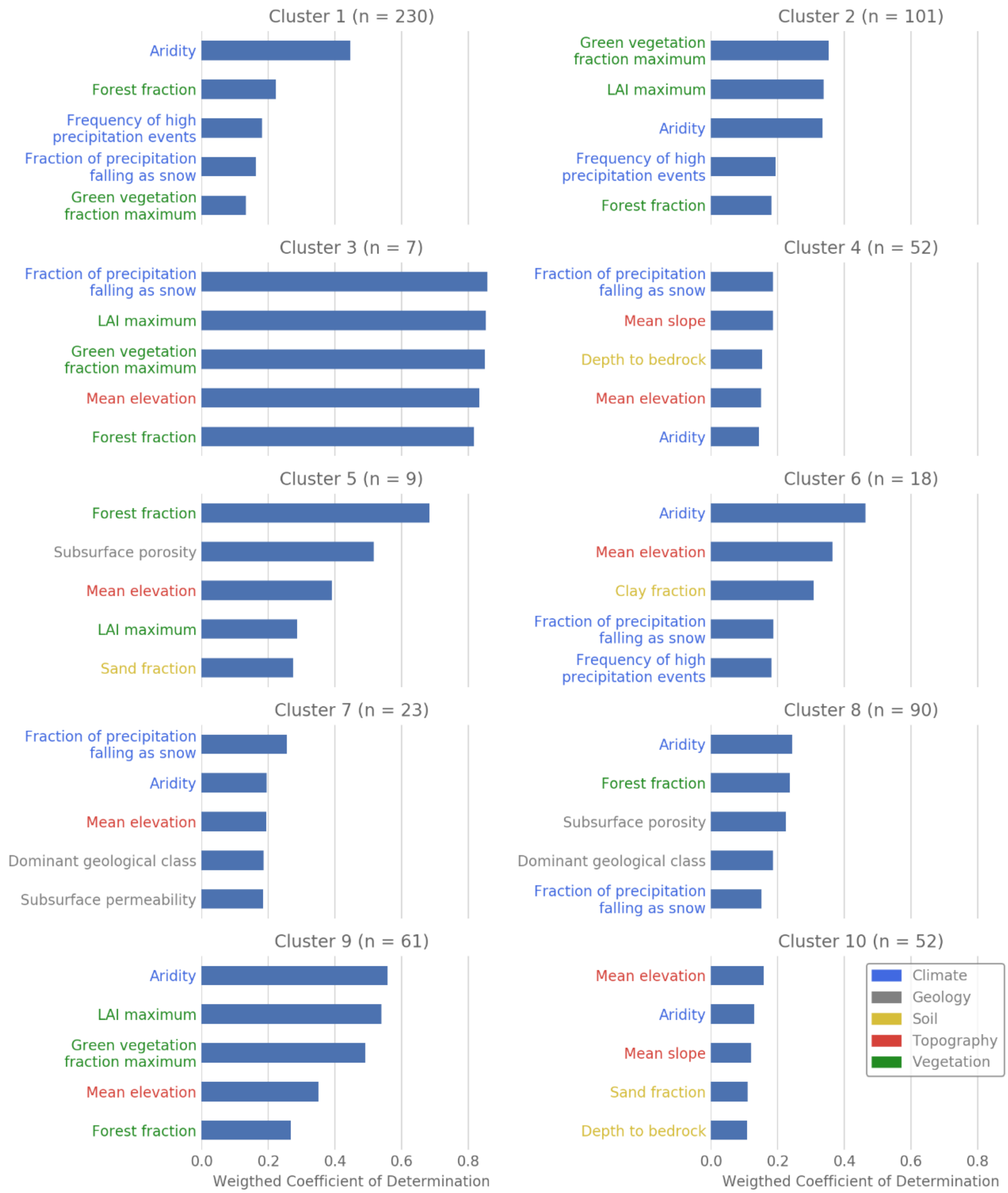


Figure 45: Importance of the catchment attributes evaluated by the quadratic regression. For the catchment clusters. Attributes colored according to their catchment attribute class.

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, 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)

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 R^2 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 R^2 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.

Overall, 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 a less strong climatic signal in those regions. 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- and catchment-attributes data cannot sufficiently explain the hydrological characteristics.

~~In addition, all catchment attributes have a high coefficient of determination in Cluster 3, while the coefficient of determination 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).~~

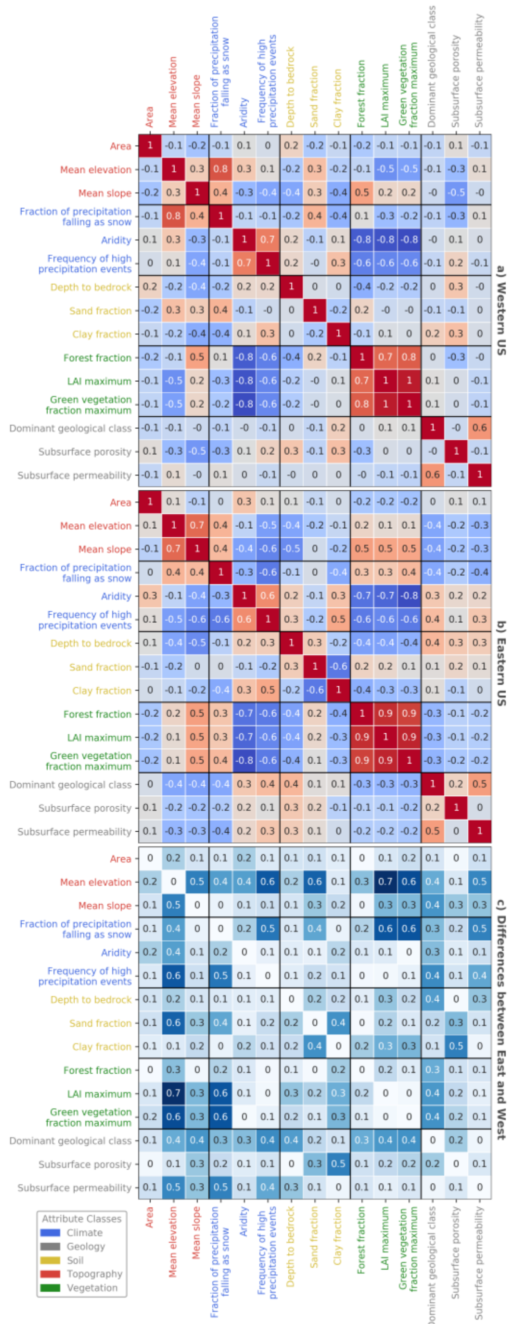


Figure 6: Correlation of all catchment attributes for western (a) and eastern (b) US and absolute differences in correlation between the eastern and western US. Eastern and western is defined by the 100th meridian.

~~This implies that climate is a good indicator for the discharge characteristics as long as the topography is homogenous. This would also explain why studies like Sawicz et al. (2011) or Berghuijs et al. (2014) were able to find strong connections between climate and their catchment clusters, as most of their catchments were located in the eastern half of the United states. This region has only few, but very distinct changes in topography such as the Apalachian Mountains and therefore climate has the largest influence. The same effect can be seen in the distribution of the clusters of this study (Figure 3). While the catchments in the eastern half of the United States form large spatial patterns of similar behaviour, the catchments in the west are patchier. This would also explain 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). Therefore clustering by climate or spatial proximity might only work in regions without abrupt changes in the topography. In addition, this is also linked to the problem that it is easier to find the most important drivers for the behavior in some regions then in others (Singh et al., 2014) and that often catchments show a surprisingly simple behavior across many different climate and landscape properties (Troch et al., 2013). The regions where it is easy to find the most important drivers show a homogenous topography, while catchments that are hard to understand with current hydrological knowledge, are controlled by a very complex interaction of factors like land use, soil or vegetation. This complex interaction is overwritten in regions with strong climatic influence.~~

~~The descriptions of the catchment clusters are summarized in Table 2. A detailed description of the clusters can be found in the appendix, together with figures showing the distribution of hydrological signatures (Figure A1) and catchment attributes (Figure A2). A list of all catchment with index, position and cluster is given in the supplementary material.~~

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 ~~coefficient of determination~~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	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 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 ~~further~~ overlay potentially 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. 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, 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 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 cluster 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). Those clusters also show distinct spatial patterns (Figure 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 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) 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 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)).

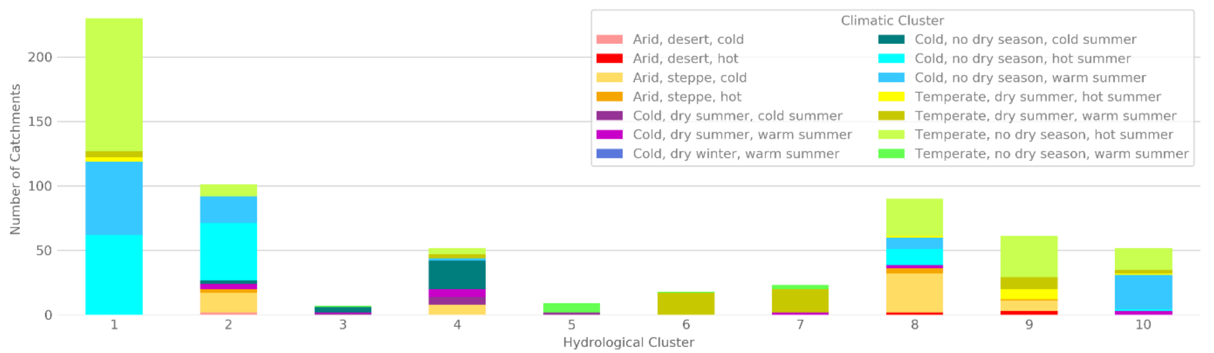


Figure 7: Membership of Koeppen-Geiger clusters (Beck et al. (2018)) in the hydrological clusters.

The picture is less clear concerning the climatic index space of Knoben et al. (2018) (Figure 8a). 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. 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, we could not identify a clear dominating catchment attribute, if we look at Cluster 4 (located in the Northwestern Forested Mountains and Florida) (Figure 5). Catchments with this hydrological behavior can be found in the space of the climatic indices of Knoben et al. 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 8b, A4) 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. This is even more remarkable, as the signatures used are linked to climate directly. For example, the signature “mean half flow date” can be seen as a measure of seasonality. 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.

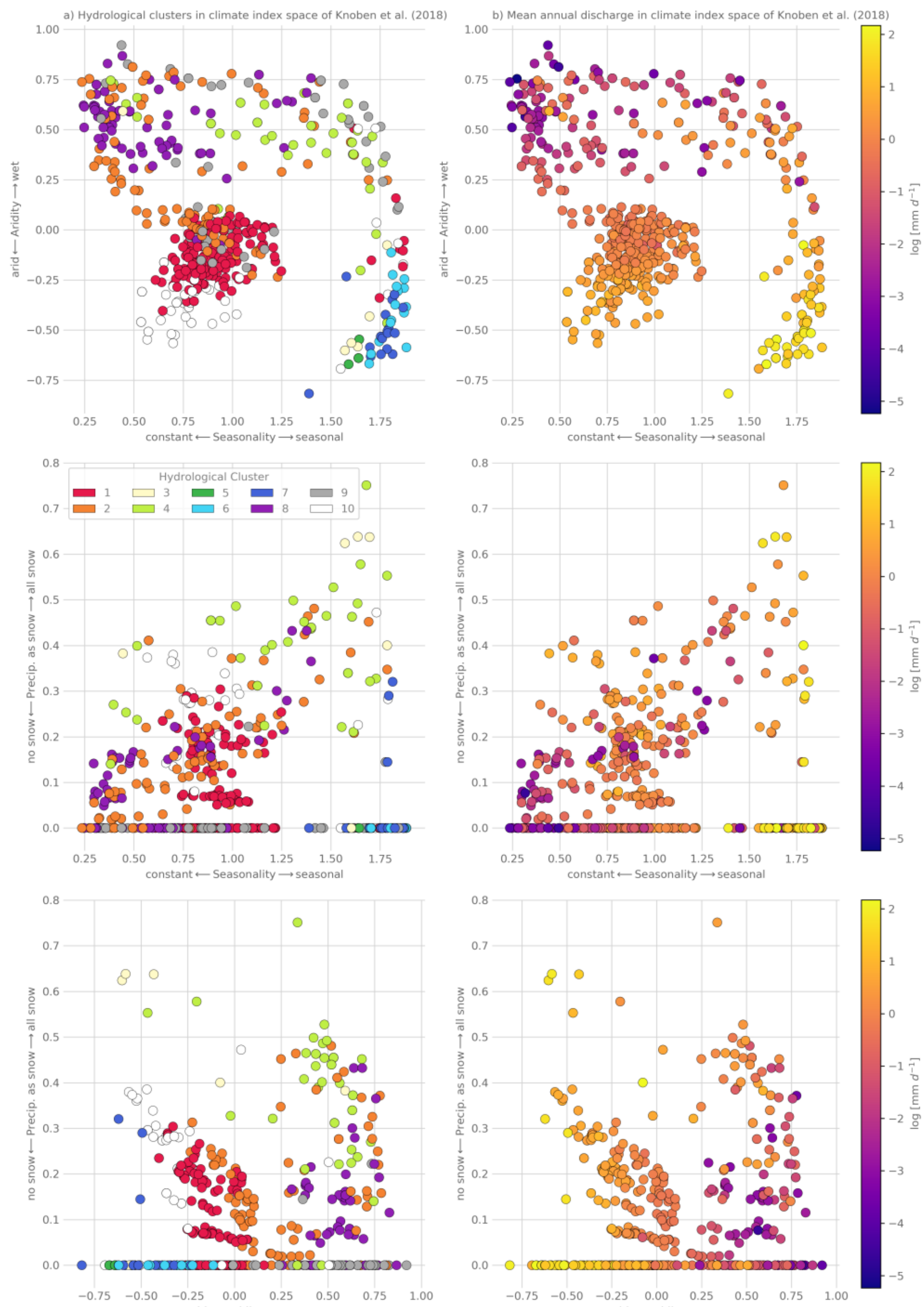


Figure 8: a) Comparison of the hydrological clustering of this study with the climate index space of Knoben et al. (2018). Single dots show the catchments and are colored by their hydrological clusters. b) Mean annual discharge for all catchments in the climate index

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 A4.

4 Summary and conclusion

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 catchment 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, 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.

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. Nevertheless, 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, it might be a worthwhile pathway to include measures of spatial variability of the climate in the single catchments. 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.

~~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 look at clusters that are derived from specific hydrological signatures. While some clusters still have aridity~~

~~as the most important factor, it can be the elevation, vegetation and amount of snow for others. We link this to the location of the catchments. The catchments that are most influenced by climate are mainly located in the eastern continental United States, where we find large regions without abrupt changes in the topography. Those catchments that are influenced mostly by other factors than aridity, show a patchier spatial pattern and are located in the western continental United States, where the topography changes on small scales. From this, we conclude that climate is the most important factor for the discharge characteristics in regions with homogenous topography. For regions with a heterogeneous topography, on the other hand, this leads to catchments that can be quite different on a very small scale, as differences in elevation and slopes create abrupt changes in most catchment attributes (e.g. soil or vegetation). This also hints why those kind of catchments are difficult to simulate . They probably have many features with a roughly equal influence on their behavior and those features alter and influence each other. This complex interaction can also lead to catchments that are quite different in their attributes, but show very similar discharge characteristics. An example for this is Cluster 4 that contains catchments from the Northwestern Forested Mountains and Florida. Two very different regions, but still the catchments show a similar behavior. This indicates that a catchment classification based only on catchment attributes is predestined to fail in regions were the main driver is not climate.~~

~~We acknowledge that the results are somewhat 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, the hydrological signatures used have been identified as the ones with clear hydrological meaning.~~

~~For further research, we think the clusters identified here can be used to explore the usefulness of the CAMELS dataset in studies dealing with parameter transferability of hydrological models, either between different types of catchment clusters or how different kinds of models perform in the same cluster. In addition, the groups of indistinct catchments should get more attention in modelling and fieldwork, as those catchments are probably also difficult to understand, because it is not clear what is causing them to behave the way they do. As long as there are catchments that cannot even be clustered by our current understanding, we as the hydrological community, still have gaps in our knowledge.~~

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).

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 ~~clustered CAMELS and analyzed the results~~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 anonymous referee~~-~~ for giving important feedback in ~~the first stages of creating this paper.~~ in the creation of this paper. We would also like to thank a~~Al~~ll 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.

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.

~~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

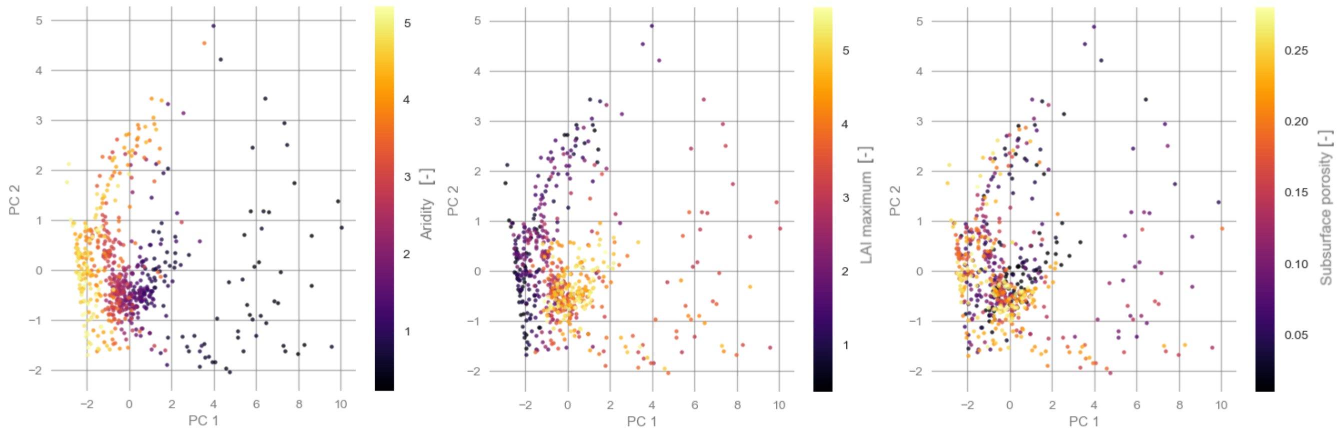


Figure A1: Patterns of catchment attributes in the PCA space of the hydrological signatures.

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.

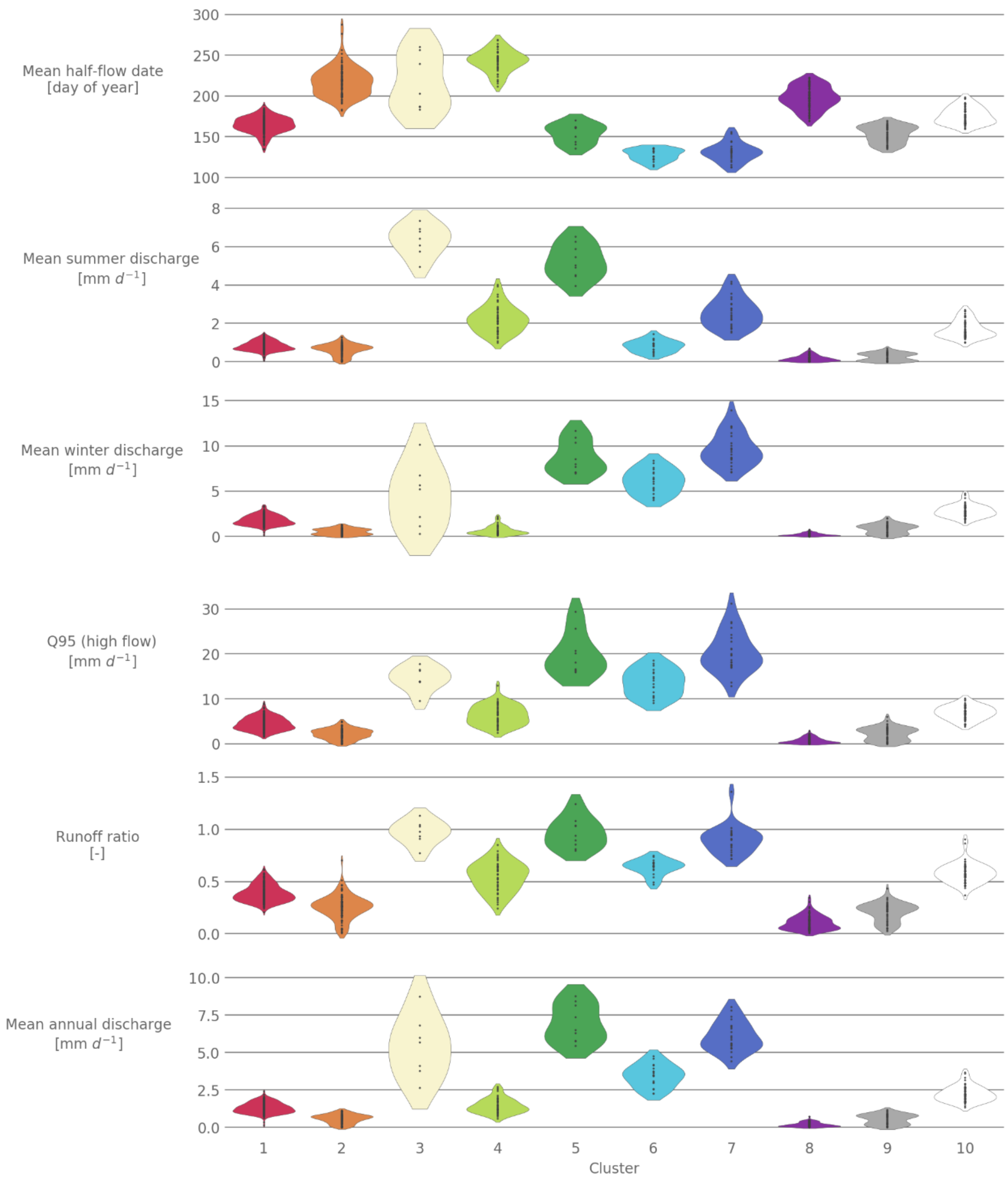
Cluster 6 catchments are also located in the Marine West Coast Forest, but cover the whole region and not only the northern part like Cluster 5. The catchments are very similar in their attributes and discharge characteristics to Cluster 5, with the exception of a lower discharge and runoff ratio. This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



695 **Figure A~~1~~2: Violin plot of the hydrological signatures sorted by catchment clusters. Single dots in the violins indicate the single catchments.**

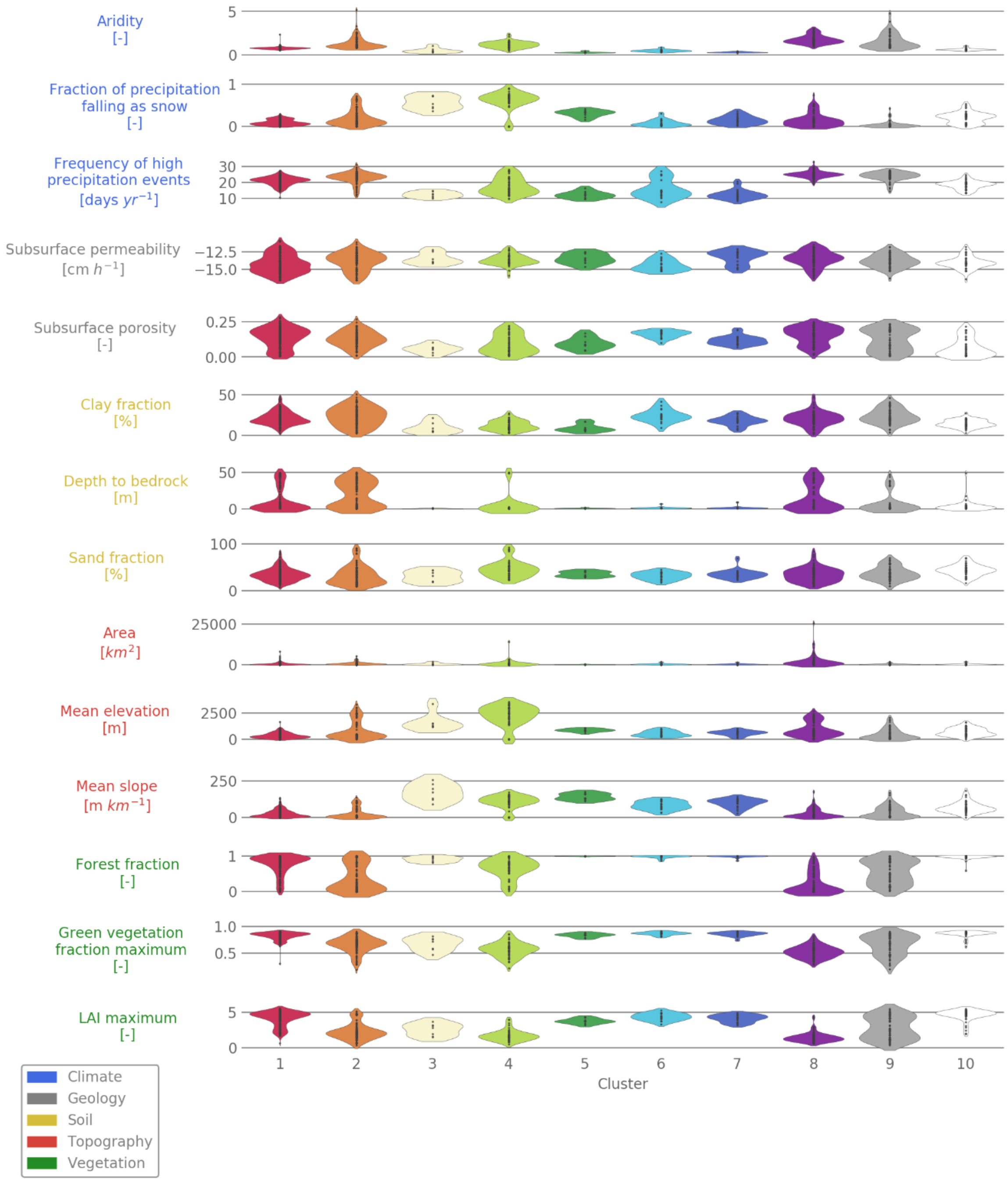
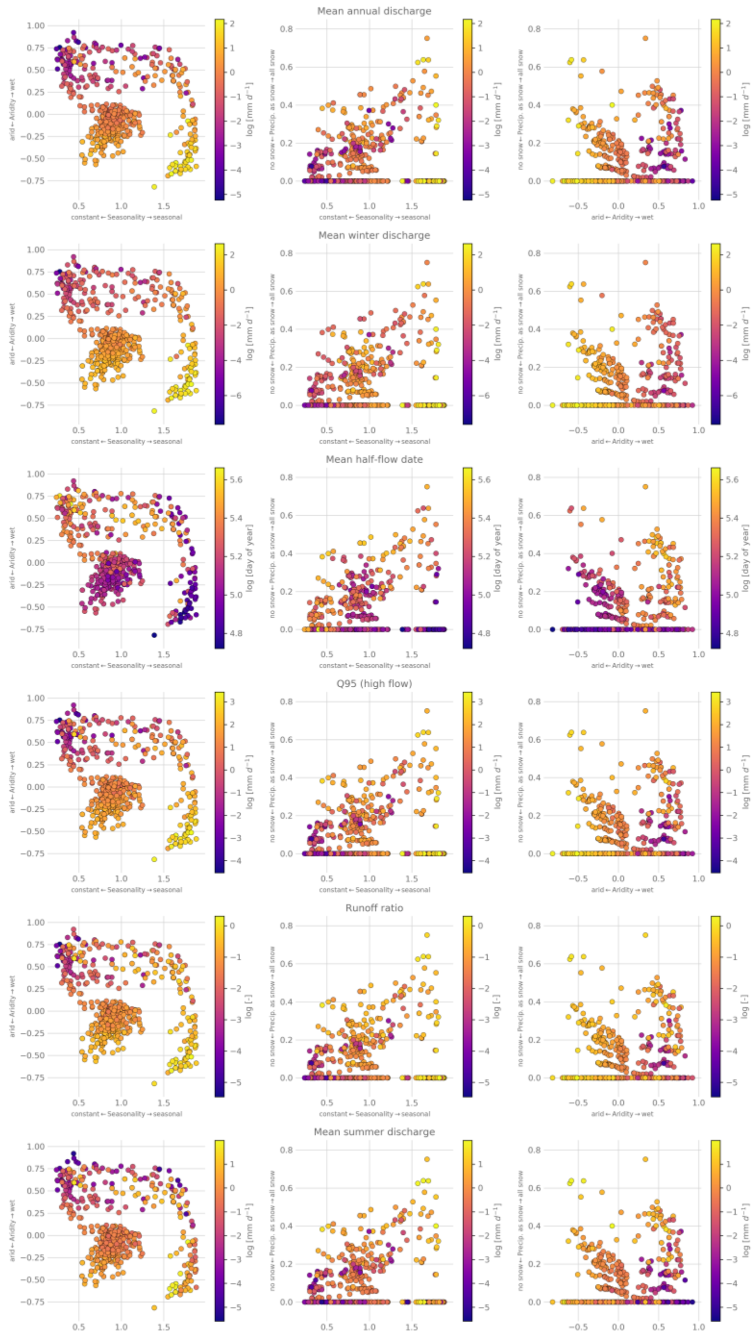


Figure A~~2~~3: Violin plots of the catchment attributes sorted by catchment clusters. Single dots in the violins indicate the single catchments.



700 **Figure A4: Hydrological signatures for all catchments in the climate index space of Knoben et al. (2018). Single dots show the catchments and are coloured 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.**