Dear editor,

Thank you very much for acknowledging the value of our work. We provide a step-by-step response to the reviewers' and commentator's comments below.
Thank you very much for considering our manuscript for publication in HESS.

On behalf of all co-authors,
Manuela Brunner

Reviewer 1

General comments:
This paper develops a functional hydrologic classification method, applies it to a large number of gages in the continental US, and assesses future changes in hydrologic regime resulting from climate change using VIC model output driven by downscaled GCMs. The paper makes an important contribution to the literature on hydrologic classification because its use of functional data analysis, in which each hydrograph is represented as a function, addresses issues associated both with top-down classifications based on catchment characteristics (lack of process connections between those characteristics and streamflow) and with bottom-up classifications based on streamflow indices (requiring selection of indices that may not fully describe the hydrologic regime). As such, the paper provides an interesting and innovative approach that is useful for assessing hydrologic regimes for climate-change impact assessment or to predict the behavior of ungaged catchments.

Specific comments:
Line 30: It should be noted upfront that hydrologic regimes, defined here as those "described by mean annual hydrographs", can encompass a broader set of variables than those used here. In particular, the functional approach that analyzes mean annual hydrographs describes the seasonal patterns of stream-flow, but not patterns that occur on either shorter timescales (e.g., flashiness) or longer timescales (e.g., interannual variability). For analyzing these types of changes in hydrologic regime, the functional analysis is less useful and particular streamflow indices must instead be used. This is important because the seasonal hydrologic regime is highly sensitive to changes in temperature in the melt region because of the snow signal, but as noted later in the paper (lines 291-292), there is no clear seasonal signal in catchments that are more sensitive to changes in precipitation to temperature. It is not that those catchments will not experience hydrologic changes due to climate change, it is just that the type of changes they experience (e.g., greater flashiness or inter annual variability) are not captured by the functional analysis and its focus on seasonal changes. Also, although at the seasonal scale of analysis meteorological variables were more significant than physiographical variables in predicting class membership (lines 228-229), it is possible that physiographical characteristics would be more significant in determining class membership when it comes to flashiness or interannual variability, because of the role of land-surface characteristics like lithology, soil, and vegetation in mediating the climatic signal.
Reply: *Thank you for highlighting this point. We agree that the functional clustering approach presented here does not consider similarities in inter-annual variability when clustering catchments, also not in the change assessment. While the approach does indeed not allow for the consideration of flashiness at an event scale, it allows for a partial consideration of flashiness as the mean annual hydrographs used for the clustering and the change analysis have a daily temporal resolution. We make the following addition when comparing the FDA to the index-based clustering approaches:*

This scheme makes better use of the seasonal and temporal information stored in the hydrological regime than index-based approaches. However, it does neither consider streamflow patterns at short, event time scales such as flashiness, nor time scales longer than a year as for instance caused by inter-annual variability.
We acknowledge in the results section that: 'Catchments without predicted regime changes may still undergo changes in specific streamflow characteristics, such as variability or low and high flows (l.274-275).'
We differentiate the discussion of class predictor strengths by adding: 'The relationship of class membership to physiographical characteristics may be weaker than the one to climatic characteristics as the clusters are formed using the mean annual hydrographs whose seasonality is strongly influenced by climate. The link to physiographical characteristics may be stronger if streamflow characteristics at an event time scale are considered.'
Modification: p.3, l. 74-76 and p.16, l.305-308

Lines 233-234: Would it be possible to include a map of nominal hydrologic class for ungaged catchments? If the random forest model can predict hydrologic class based on meteorological and physiographical variables, it should be able to apply the classification and predict hydrologic class for all the catchments in the CONUS (assuming data for the predictor variables are available CONUS-wide). That would be an interesting map to see because it would further illustrate the spatial contiguity and extent of the hydrologic classes beyond the gaged catchments in Figure 2.
Reply: *We agree that it would be nice to produce a map of predicted regime classes over the whole CONUS. We did not do this for two main reasons: First, the dataset of physiographical variables we used for this analysis is limited to the 671 catchments in the CAMELS dataset* [*Newman et al.*, 2015; *Addor et al.*, 2017]. *Furthermore, our classification is limited to streamflow regimes resulting from natural conditions. A classification for all catchments in the CONUS would need to encompass classes for 'human-influenced' catchments where the streamflow regime has been (strongly) altered by water abstractions and transfers, reservoir operation or other human interventions.*

Technical corrections:
Line 243: Typo: "Klomogorov" should be "Kolmogorov".
Reply: *We corrected this typo.*
Modification: p.12, l.248

Line298: Is "temporally" meant to be "temporarily"? That makes more sense to me
Reply: *Yes, we changed temporally to temporarily.*
Modification: p.19, l. 343

Reviewer 2 : Florian Ulrich Jehn

General evaluation:
This paper proposes a new method to cluster catchments based on the temporal information in their hydrological regime and uses the found regime clusters to evaluate how climate change will change the regimes clusters. I think this is an interesting approach and yields good results, especially the changing of the regimes due to climate change. In addition, the paper is overall well written and has a good flow to it. I think it can be published after minor revisions. However, I have one larger points where I think clarification is necessary.

Main point:

- Line 98: Why those five spline basis functions? How can you be sure that those are enough to represent diverse regimes? It there a connection between using five spline basis function and finding five streamflow regimes? I think this part should be extended to make it clearer what those decisions where based upon.
Reply: *Thank you for expressing your concern regarding a suitable choice of spline basis functions. We chose five spline basis functions because our tests showed that a further increase in the number of spline bases did not further improve the clustering results. This is confirmed by the overall silhouette width, which is for more spline basis functions (6 to 10) lower or very similar to the one for five basis functions (at five clusters). There is no relation between the number of spline bases and the number of regimes chosen. A choice of four or five clusters would also be optimal if more than five spline bases were used.*
We added the following specification to the text: 'The suitability of five spline basis functions is confirmed by the overall silhouette width, which is for more spline basis functions (6 to 10) lower or very similar to the one for five basis functions.'
Modification: p.4, l.106-107

Minor Points:
- Line 47: "The use of catchment characteristics can be problematic because there is often no clear link between these characteristics and streamflow indices (Ali et al., 2012; Addor et al., 2018)." I think this is worded a bit too strict. For example, Addor et al. 2018 indeed showed that there are differences between the link of catchment characteristics and streamflow indices, but they also showed this connection can be relatively strong for some catchment attributes. Overall, I think this section should be less dismissive of the findings of the cited papers.
Reply: *We did by no means intend to be dismissive of findings of other papers as we appreciate their work. Furthermore, we do not think that it is a bad thing to find weak relationships between certain streamflow and catchment characteristics. We rephrased the sentence to: 'The use of catchment characteristics is not always beneficial as certain streamflow indices do not show clear links to these characteristics* [*Ali et al.*, 2012; *Addor et al.*, 2018]*'. Furthermore, based on the point raised by the other reviewer, we stress that the relation between catchment characteristics and streamflow might be more apparent for event-based signatures.*
Modification: p.2, l.47-48

- Figure 2 and its discussion: As you are already citing my paper, I hope it is appropriate to mention that that the final version is now published in HESS (https://www.hydrolearth-syst-sci.net/24/1081/2020/hess-24-1081-2020.html) and discusses the different flow regimes in CAMELS in more depth than before (Figure 6). I think the results are very similar to the ones in this paper, but also show that the flow regimes found here can be split in more distinct groups. However, this might be more of a question of the desired granularity.
Reply: *Thank you for pointing out the necessity to update the reference. We point out the similarity between the clusters resulting from your and our analysis by saying: 'The five regime clusters identified also show spatial similarities with the ten catchment clusters formed by (Jehn2019) for the same set of catchments using a small set of hydrological streamflow characteristics.' (l.225-226). For some applications it may indeed be useful to have more distinctive clusters.*

- Line 85: How is satisfactory model performance defined?
Reply: *Melsen et al.* (2018) *who produced the simulated streamflow data used in this study provide simulations only for the subset of 605 catchments. They chose to run simulations for this subset only as different data sources disagreed on catchment size for the remaining catchments.*

A reliable estimate of catchment size is crucial to ensure accurate forcing input for the lumped model. We reformulated the sentence as follows: 'In contrast, the regime change analysis uses streamflow simulated by the hydrological Variable Infiltration Capacity (VIC) model for a subset of 605 catchments, for which reliable data on catchment area was available at the time the simulations were produced [*Melsen et al.*, 2018]. *Kling-Gupta efficiencies obtained over these basins with VIC varied from a first quartile of 0.47, a median of 0.6 and a third quartile of 0.71, with the lowest values obtained in the Great Plains'. (l.145, p.6)*
Modification: p.4, l.91-94

- Line 90 and following, code availability: I did not see any link to a code repository for this paper (my apologies if I missed it). I think in a paper that does propose a new method, it is important to provide the code used. While the method section explains the idea well, it is still a non-trivial task to recreate the method of this paper without any code examples to work with.

Reply: *It is correct that we have not yet provided a link to the repository with the catchment clusters. This is because we wanted to wait for the DOI of this manuscript to be available before we created a DOI for the dataset. The dataset can now be accessed via HydroShare: The link to the dataset was added to the manuscript (*https://doi.org/10.4211/hs.069f552f96ef4e638f4bec281c5016ad).
To facilitate the reproduction of the functional data clustering approach, we added details on the R-packages and functions we used: 'The analysis is performed in R using the packages fda.usc [*Febrero-Bande and Oviedo de la Fuente*, 2012] *and fda* [*Ramsay et al.*, 2014] *and the following functions: (1) conversion of regimes to functional data objects: fdata, (2) creating of B-spline basis functions: create.bspline.basis, (3) computation of spline coefficients for all regimes: Data2fd. The clustering into regime classes is performed using the R-package stats (R Core Team, 2019). A Euclidean distance matrix is computed using the matrix of n= 671×5 spline coefficients (Figure 1.2 a–b) (dist). We use a hierarchical clustering algorithm (hclust) allowing for non-elliptical clusters (Gordon, 1999) with Ward's minimum variance criterion, which minimizes the total within-cluster variance (Ward, 1963). To identify an optimal number of clusters, we cut the tree a tk= 2,...,30 clusters (cutree) and compute the mean silhouette width (Rousseeuw, 1987), which provides a measure of clustering validity, for the different numbers of clusters'*
Modification: p.9, l.219 and p.4, l.108-116

- Line 226: I am unsure if avoiding small clusters is a sign of a good clustering. As river behavior is a natural process, I would expect it to follow some kind of normal distribution, which would result in some bigger clusters and some smaller, more extreme clusters. For future research, it might be a good idea to explore a continuous classification as done by Knoben et al (https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018WR022913) or a fuzzy clustering approach to avoid the arbitrary cut-off points of clusters.

Reply: *One can always form new clusters of 'special cases' to decrease the within cluster variability. However, the formation of more clusters in our case resulted in a decrease in average silhouette width, which is not desirable. Furthermore, the formation of more clusters with smaller in-between cluster differences, would have diverted the focus from the detection of major regime changes to the detection of minor regime changes. We removed the sentence suggesting that fewer clusters were better than more clusters. We agree that the use of probabilistic instead of deterministic class memberships may be desirable for some applications. For this particular application, however, we found deterministic clusters to be more appropriate.*
Modification: p.17, l.301-302

- Line 256: The difference here might again be a question of granularity. Especially the

catchments in Florida behave very uniquely.
Reply: *The finding that the correct regime class is not simulated in certain regions when forcing the model with GCM output is likely related to the fact that certain processes in these areas are not well represented by GCMs.*

- Figure 5: It is very difficult to distinguish the lines from each other here. I think it might be a good idea to increase the size of this figure. Also, I would recommend to use more easily distinguishable colors.
Reply: *Thank you for pointing out the need to increase the legibility of this figure. We rearranged the subplots of this figure in order to increase the size of the individual subplots and darkened the color of the control regime to increase contrast with respect to the regimes simulated using the GCM output.*
Modification: Figure 5

Reviewer 3: Genevieve Ali

COMMENTS
In this manuscript entitled "Future streamflow regime changes in the United States: assessment using functional classification", two main goals are pursued: (1) develop a catchment classification scheme for streamflow regimes, and (2) use this scheme to evaluate changes in future flow regimes. Contrary to the majority of previously published catchment classification efforts, here the authors decided not to rely on stream-flow indices. Instead, they are using a functional approach via which the shapes of mean annual hydrographs are classified, this in order to retain temporal autocorrelation information. Overall, the manuscript is of appropriate length, well written and with good-quality figures and tables. The dual focus of the manuscript on catchment classification and climate change impact assessment is very interesting, and I agree with the authors about their description of the advantages of functional classification. I did find that a few statements made in the manuscript warranted clarification, and that some details regarding the datasets, process interpretations, or linkages with existing literature were lacking (see specific comments below). With revisions, I believe that this manuscript will be interesting to the HESS readership, and a great addition to our growing body of literature on catchment classification.

Reply: *Thank you very much for your thorough review. We provided missing details on the dataset, added a few clarifications, and extended the discussion as suggested. The changes are discussed in detail below.*

SPECIFIC COMMENTS

N.B.: page and line numbers are noted as PX (page X) and LX (line X).

Section 2.1: Given the international readership of HESS, I think that more detailed in-formation is needed about the catchment selection criteria. For people not familiar with the CAMELS dataset, it is quite unclear what is meant by "minimum human impact": is the human impact assessed in terms of catchment-wide land use (that would mean no agricultural or urban catchment), or river regulation? And how may the answer to that question affect the generalization potential of the manuscript conclusions? In other words, the authors should discuss the limitations associated with not considering human-impacted catchments in the present study... Also, how was the 1981-2018 data period chosen for the analysis?

Reply: *Thank you for pointing out the need for clarification. We focused on catchments with minimum human impact to be able to look at the effects of climate change on flow regimes in isolation. The catchments belong to the HCDN-2009 network* [Lins, 2012]*, which is a set of stations deemed suitable for analyzing hydrologic variations and trends in a climatic context. The dataset consists of catchments with natural flow conditions undisturbed by artificial diversions, storage, and other activities in the drainage basin or the stream channel and show less than 5% imperviousness as measured by the National Land Cover Database* [*Jin et al.*, 2013]*. We added this information to the text. We also specified that the period 1981-2018 was chosen 'as data for this period was available for most stations in the dataset'.*
As suggested, we added a discussion on the limitations of the classes in the case human-impacted catchments are of interest: 'The streamflow regime classes identified here do not comprise classes of catchments with major flow alterations as the clustering was performed using streamflow regimes from catchments with minimal human impact. The five classes proposed here are therefore of limited use if a problem requires including catchments with strong human flow alterations. A flow regime of a regulated stream may still be attributed to one of the five regime classes identified if the altered regime shows similarities with the flow seasonality and variability of one of the 'natural' classes. However, if flow alteration leads to the emergence of regimes clearly distinct from those observed under natural conditions, additional regime classes would be necessary. In addition, the relationships between catchment characteristics and class memberships would need to be revised to enable the assignment of ungauged catchments to one of the classes in the updated set.'
Modification: p.3, l.85-89 and p.18, l.314-321

P5 L120: There is a reference to characteristics with missing values. Which characteristics (or types of characteristics) are the ones with missing values? Did omitting them lead to biased results?

Reply: *Among the 33 characteristics available, 2 had missing values (i.e. 'second most common geologic class in the catchment' and 'subsurface porosity'). They both belong to the class of geological characteristics comprising 7 characteristics in total, which means that we were still able to consider 5 characteristics related to geology. We found, however, that these geological characteristics were of minor importance for explaining regime class membership. We specified the two classes with missing values in the manuscript.*
Modification: p.5, l.132

Section 2.5, specifically L180-182: How was the comparison made, exactly, from a quantitative or statistical standpoint? Using contingency tables or crosstabs? Or something else? This is a bit unclear to me.... Maybe because I was expecting a statistical comparison when in fact, it is not what was done...

Reply: *We checked whether the predicted future class corresponded to the class of the reference simulation. The outcome of this check is binary: 0: predicted future class corresponds to reference class, 1: predicted future class differs from reference class. The results of this comparison are shown in Figure 6 (bars on the left). For the catchments with regime changes, we then identified the direction of change using a contingency table of counts (Figure 6, colored bars on the right).*

We specified in the text that: 'We then compare the predicted future classes to the class of the corresponding reference simulation using a contingency table of counts.'
Modification: p.7, l.195-196s

Figure 3: The different (graphical) features of the boxplots should probably be de-scribed in the figure caption. I assume that the horizontal black lines refer to the medians.... what do the whiskers represent, though: 1 interquartile range (IQR), 1.5 IQR, min and max values, or something else? Are there no statistical outliers associated with each cluster, i.e., each individual box?

Reply: *Thank you for pointing out the need for specification. We added the following text to the caption: 'The black lines in the boxplot indicate the median, the upper and lower whiskers correspond to 1.5 * R_{IQ}, where R_{IQ} is the inter-quartile range. Outliers are not displayed.'*
Modification: p.10, caption Figure 3

P9 L203-204: That should not be a surprise, given that the flood and drought definitions are hydrograph-based.... or am I missing something?

Reply: *The droughts and floods were determined using a threshold-level and a peak-over-threshold approach, respectively while the regime classes was determined using the mean annual hydrographs where extremes are smoothed out. But yes, we would expect some correspondence between the streamflow regime of a catchment and the types of extreme events it experiences. We here show that flood and drought event characteristics of the different streamflow regime classes are indeed distinct (Figure 3 in the manuscript).*

P9-10, L207-210: The text description, here, does not underline that strong of a contrast between the weak winter regime and the strong winter regime. Maybe it can be rephrased for the contrast to be expressed more strongly?

Reply: *We added the following sentence highlighting the differences between catchments with a weak and strong winter regime: 'Compared to catchments with a weak winter regime, catchments with a strong winter regime lie at higher elevations, show higher fractions of snow and are characterized by larger flood magnitudes.'*
Modification: p.11, l. 226-228

P10 L217-218: That would explain why there is such a large degree of spatial contiguity/spatial autocorrelation within each cluster. However, it is a bit unclear to me, from the text, whether a RF classification using climatological variables only performs equally as well as – or better than – a RF classification that used both climatological and physiographic variables.

Reply: *You are right, we did not discuss whether the random forest model profits from including additional physiographical characteristics in addition to climatological ones. We just discussed variable importance in the context of the 'full model' including all potential explanatory variables. If physiographical variables are excluded from the random forest model, the prediction error increases from 10% to 12%, which corresponds to a marginal decrease in model performance. So yes, a random forest classification using climatological variables performs almost equally as well as a model also including physiographical variables. We added the following sentence to the text: 'Excluding these physiographical explanatory variables from the random forest model results in only a small decrease in prediction performance (prediction error*

12%).'

Modification: p.11, l.236-238

P10 227: The authors stated that "However, our clustering scheme avoids the formation of very small clusters seen in Jehn et al. (2019)." First, what might explain this? Second, the authors seem to imply that having very small clusters is an inconvenient, and I am not sure I agree – very small clusters could represent very local conditions or hotspots, which are real. The authors should either rephrase or at least nuance their statement to clarify what they mean.

Reply: *We did not intend to suggest that forming small clusters is necessarily a bad thing and therefore reformulated the sentence using neutral wording: 'However, our clustering scheme results in larger clusters than the ones seen in Jehn et al. (2019).' This is mostly related to the fact that we chose to work with fewer clusters. If we further increased the number of clusters to e.g. 7 instead of 5 clusters (Figure 1 in this response to the reviewer), we would also introduce very small clusters. We would further split up the melt-regime cluster and the New-Year's-regime clusters. This does, however, not further improve cluster distinctiveness as measured by the mean silhouette width.*

Modification: p.17, l.301-302

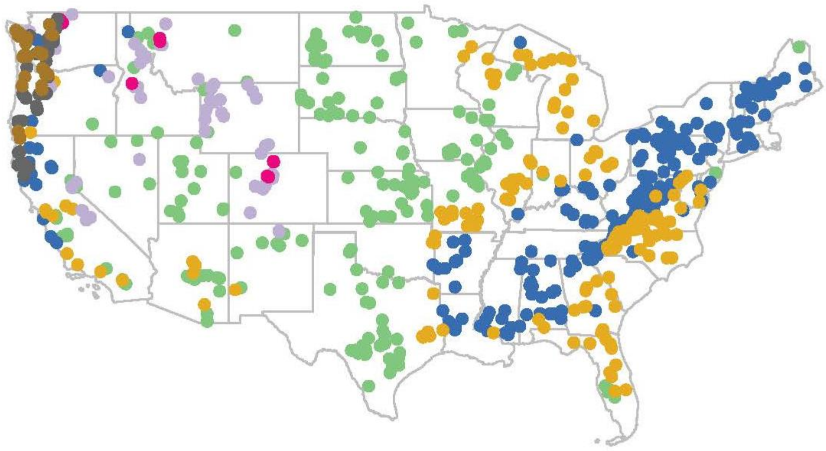

Figure 1: Map of 671 catchments in the dataset clustered into 7 streamflow regime classes. Each color represents a different class.

P10 L230-234: The authors wrote that "The strong link between regime classes and meteorological and physiographical catchment characteristics allows for the attribution of ungauged catchments, where streamflow data are not available, to one of the regime classes, which is potentially very useful for the prediction of streamflow characteristics in ungauged basins". I am not sure where that statement is coming from, as ungauged catchments were not examined in the present study. I agree that the present study might have interesting implications for predictions in ungauged catchments, but this statement, as written, reads as a result when in fact it is an interpretation. In the same line of thought, I wonder whether it would be possible to have separate Results and Discussion sections in the manuscript. There are a few instances, in the text, where it can be tricky to distinguish whether a plain result/fact is being stated, or whether a hypothesis/interpretation is being put forward.

Reply: *We split up the Results and Discussion section into two sections in to more clearly distinguish between the results of our study and their implications. It is correct that the focus of our study is not on prediction in ungauged basins. However, we show that a random forest model* *fitted to climatological and physiographical characteristics is well able to attribute a catchment to one of the regime classes without having any information on streamflow (class prediction error 10%, see l. 118-124 and l.215-119). We add the following sentence to the methods section: 'To further investigate the physiographical and climatological controls on regime class membership and to check whether regime classes can potentially be predicted for ungauged catchments, we perform a random forest classification'. Thanks to its low prediction error, this random forest model enables attributing of ungauged catchments to one of the regime classes. As we do not go into detail on this aspect, we moved the statement to the new Discussion section and clarified it as follows: 'The strong link between regime classes and meteorological and physiographical catchment characteristics enables attributing ungauged catchments, where streamflow data are not available, to one of the regime classes. This attribution can be achieved by using the first random forest model fitted in this analysis enabling predictions of regime class membership using physiographical and climatological characteristics. The ability to attribute of an ungauged catchment to one of the regime classes is potentially very useful to predict of streamflow characteristics in ungauged basins.'*

Modification: Section 4 Discussion, p.5, l.128-129, p.18, l.310-312

Figure 4: This figure is quite interesting but the comparison of "climate sensitivity" between observations and simulations appears quite qualitative. I wonder: 1) How were the five example catchments showcased in this figure chosen (or, are those sites representative of median cluster conditions)?; and 2) Was a quantitative method of comparison between observations and simulations used for all catchments?

Reply: *We chose one regime per cluster and the sites do not necessarily represent median cluster conditions. Yes, we also applied a quantitative method to evaluate 'climate sensitivity' over all catchments (l.241-243). We added that: 'The sensitivity gradients are computed on the response surface of each catchment in the horizontal direction for temperature and in the vertical direction for precipitation.' The results of this quantitative evaluation are summarized in Figure 2 in this response to the reviewers. The statement: 'Higher mean precipitation leads to higher mean discharge independent of the catchment and regime. The reaction of streamflow to temperature, however, seems to depend on the catchment because the relationship between mean temperature and mean discharge is generally weak and can be positive or negative. (l.238-240)' can therefore be generalized to the entire dataset. We preferred to show the sensitivity grids for a few catchments as we think that these examples nicely illustrate the mechanisms we see for the whole dataset.*

Modification: p.12, l.250-251

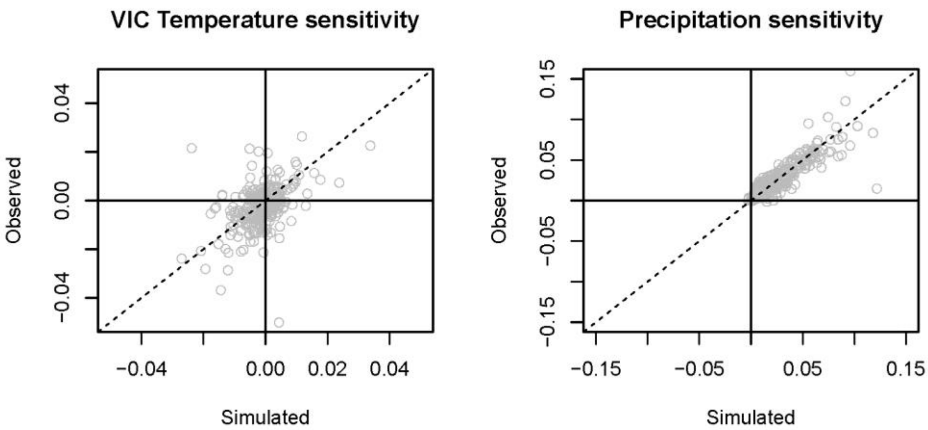

Figure 2: Observed vs. simulated sensitivity gradients for temperature (left) and precipitation (right) over all catchments computed using climate sensitivity grids as displayed in Figure 4 of the manuscript for five example catchments.

P11 L240: The authors refer to a "visual analysis"; were all plots for all 605 catchments visually analyzed?

Reply: *We computed such sensitivity grids for each catchment and used them to compute horizontal and vertical sensitivity gradient as outlined in the response to the previous question.*

P11 L243-244: The Methods section should explicitly state what the Kolmogorov-Smirnov test was used for, the assumptions being it, and the null and alternate hypotheses (so that readers know what the test results mean). Also, a test cannot be rejected: we can only reject or fail to reject a null hypothesis, so that sentence should be reworded.

Reply: *We rephrased the sentence to: '(Kolmogorov-Smirnov test does not reject the null hypothesis that observed and simulated gradients were drawn from the same continuous distribution at level of significance alpha=0.05.)'*
Modification: p.12, l.248-251

Figure 5: Lines are a bit difficult to distinguish on this figure; making it larger and changing the symbology might help.

Reply: *Thank you for your suggestion. We rearranged the plot into a 3 rows, 2 columns format to increase the size of the individual subplots. In addition, we darkened the color of the control regime to increase contrast with respect to the regimes simulated using the GCM output.*
Modification: Figure 5

P12 L258-259: The authors wrote "In contrast, regimes with a strong seasonality such as strong winter and New Year's regimes are well simulated". What about the melt regime, which is also highly seasonal?

Reply: *This statement is also valid for melt regimes and we added this regime type to the list.*
Modification: p.15, l.265

Figure 7: If the black circles mean no regime change, the legend should state so.
Reply: *Yes, black circles refer to no regime changes. We added this to the legend of Figure 7.*
Modification: Figure 7

COMMENTS SPECIFIC TO DISCUSSION ELEMENTS WORTH INCLUDING IN THE MANUSCRIPT

Discussion comment #1: In the present study, regime clusters appear equivalent to clusters derived based on physiographic similarity and clusters derived based on climatological similarity... this is contrary to studies published by Ali et al. (2012) and Oudin et al. (2010) – in a comforting way, I might add – and this should probably be discussed. The "overlap" or agreement between the different classifications bodes well for using climatic and physiographic information as a proxy for streamflow regime types. The fact that an agreement was found in the present study and not in others may be due to the fact that here, functional data were used instead of select streamflow indices.

Reply: *Thank you for suggesting to expand the discussion on this aspect. We added the following discussion point: 'We find functional data clustering to be a useful tool for identifying clusters of catchments with not only similar streamflow regimes but also similar catchment, meteorological, flood and drought characteristics. This similarity corroborates findings by Bower et al.* (2004) *and McCabe and Wolock* (2014) *who established a clear link between similarity in streamflow seasonality and climatic and physical similarity. However, it is in contrast to findings by Ali et al.* (2012) *who found that catchments similar with respect to a set of flow indices are not necessarily physically similar. Explicitly including seasonality or information on the temporal autocorrelation of regimes may therefore help to identify clusters of catchments which are not only hydrologically but also physically similar.' A reference to Oudin et al. (2010) was added to the introduction.*
Modification: p.17, l.290-295, p.2, l. 48

Discussion comment #2: It is not a study limitation per se, but the authors may want to discuss the rationale for using functional streamflow data classification (to preserve temporal information) while NOT using climate time series (e.g., mean annual hyetograph) for classification purposes. When I started reading the manuscript, I was puzzled by the fact that a classification based on temporally autocorrelated data (i.e., whole annual hydrographs) was going to be compared to a classification based on climate indices. In other words, I wondered how the analyses would turn out given that different regions may have similar values of mean annual precipitation, even though the temporal distribution of that precipitation may be skewed in some places but not elsewhere. In the end, the authors found that they could neglect the temporal information included in climate time series and still manage to use that climate information (i.e., the climate index class) as a good proxy for streamflow regime class (which, itself, is based on temporally autocorrelated data). That warrants discussion, I think, as it is a bit counter-intuitive (to me, anyway...)

Reply: *Our functional streamflow regime clustering approach is indeed solely based on the mean annual hydrographs and the temporal autocorrelation contained therein. It does not rely on climate time series. The information on climate characteristics is only used to see whether the hydrological regime clusters are also climatologically meaningful. We clarify this in the introduction by saying: 'This scheme makes better use of the seasonal and temporal information stored in the hydrological regime than index-based approaches and is solely based on streamflow information (i.e. no climatological information is used).' We indeed find that these clusters formed according to mean annual hydrographs are distinct in terms of climate and physiographical characteristics (Figure 3 in the manuscript). The good predictive power of a random forest model in correctly attributing catchments to a regime cluster based on climate and physiographical characteristics supports this (l.215-217).*
Modification: p.3, l.74-76

Discussion comment #3: The authors may want to use the concepts of resistance, resilience and synchronicity discussed by Carey et al. (2010): those concepts partly echo what the authors are referring to as "climate sensitivity".

Reply: *Thank you for this suggestion. We extend the introduction to the climate sensitivity analysis as follows: 'In the climate sensitivity analysis, we assess whether the hydrological model reacts to changes in mean temperature and precipitation in the same way as observations. In terms of precipitation, this corresponds to checking whether the model captures the resistance of a catchment, i.e. the degree to which runoff is coupled with precipitation Carey et al.* (2010)*.'*
Modification: p.7, l.170-172

EDITORIAL SUGGESTIONS
P2 L30: "illustrate the hydrological functioning" seems more appropriate than "govern the hydrological functioning", since the authors are referring to streamflow regimes.
P2 L31: I think that the phrase "influencing streamflow variability" should be changed.... Otherwise the whole sentence read as "The characteristics of streamflow regimes [influence] streamflow variability and seasonality", which reads as a circular statement.

Reply: *We rephrased this sentence to: 'The characteristics of streamflow regimes, as described here by mean annual hydrographs, include streamflow variability and seasonality and influence the hydrological functioning of a catchment.'*
Modification: p.2, l. 30-31

P10 L217: "shows that the the most important variables for" SHOULD BE CHANGED FOR "shows that the most important variables for"
Reply: *We eliminated the duplicate 'the'.*
Modification: p.11, l.234

P11 L243: "Klomogorov–Smirnov" SHOULD BE CHANGED FOR "Kolmogorov-Smirnov"
Reply: *We fixed this typo.*
Modification: p.12, l.248

P13 L274: "In contract" SHOULD BE CHANGED FOR "In contrast"
Reply: *We fixed this typo.*
Modification: p.15, l.265

REFERENCES CITED IN THIS REVIEW

- Ali, G., Tetzlaff, D., Soulsby, C., McDonnell, J. J., and Capell, R. (2012), 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
- Carey, S.K., Tetzlaff, D., Seibert, J., Soulsby, C., Buttle, J., Laudon, H., McDon-nell, J., McGuire, K., Caissie, D., Shanley, J., Kennedy, M., Devito, K. and Pomeroy,J.W. (2010), Intercomparison of hydroclimatic regimes across northern catchments: synchronicity, resistance and resilience. Hydrological Processes, 24: 3591-3602. doi:10.1002/hyp.7880
- Oudin, L., Kay, A., Andréassian, V., and Perrin, C. (2010), Are seemingly physically similar catchments truly hydrologically similar? Water Resources Research, 46, W11558,doi:10.1029/2009WR008887

Commentator: Wouter Berghuijs

This paper presents an interesting analysis of streamflow and its changes across the USA. The paper i) classifies catchments based on their existing flow regime, and ii) assesses how these different flow regimes are expected to change into the future. Overall this paper seems like an interesting and relevant contribution to HESS, and enjoyed reading the paper. This short comment is not intended as a full review of the paper, but I hope that sharing the below thoughts may help to strengthen the paper.**I am aware that there is some degree of (what may be classified as) self-advertising in this comment, but my comments can be addressed without (again) citing the single self-reference that I provide.**
A key assertion and motivation of this study is that hydrological classifications have not really incorporated "temporal information in clustering hydrological catchments, [...] even though such information is potentially very useful". This statement is used as a motivation to develop a classification that incorporates such information. Overall, this is a good idea. However, while many studies indeed ignore the temporal aspect, there are existing studies that explicitly incorporate this information into their classifications (leading to very similar types of classifications as presented in this paper). Grouping based on seasonal regimes have been introduced a long time ago: e.g.:
Pardé, M. (1960). The river regime in New Zealand. Revue de Géographie Alpine , 48(3), 383-429. (And probably also in earlier works of Pardé and others) and have been applied globally: Haines, A. T., Finlayson, B. L., & McMahon, T. A.(1988). A global classification of river regimes. Applied Geography, 8(4), 255–272.https://doi. org/10.1016/0143-6228(88)90035-5, which classified global streamflow regimes into very similar type classes as done in the presented HESSD manuscript. (However, obviously with a greater variety of classes since the global spectrum of river flows was taken into account). A global analysis has been updated: Knoben, W. J., Woods, R. A., & Freer, J. E.(2018). A quantitative hydrological climate classification evaluated with independent streamflow data. Water Resources Research, 54(7), 5088-5109. This Knoben study also includes temporal information of streamflow classes into the final classification, which again is very similar in nature to what is presented in the presented HESSD manuscript. Different to the presented manuscript is that the classification metrics are not based on streamflow themselves directly. However, the classes it produces are shown to have very similar within-class seasonal streamflow regimes, which seems to make them functionally equivalent to what is presented in the HESSD paper. In addition, such analyses are also available for the United States, focusing on seasonal streamflow regimes: Coopersmith, E., Yaeger, M. A., Ye, S., Cheng, L., & Sivapalan, M.(2012). Exploring the physical controls of regional patterns of flow duration curves-Part3: A catchment classification system based on regime curve indicators. Hydrology and Earth System Sciences, 16(11), 4467.and seasonal streamflow (and all other water components) regimes: Berghuijs, W. R.,Sivapalan, M., Woods, R. A., & Savenije, H. H. G. (2014). Patterns of similarity of seasonal water balances: A window into streamflow variability over a range of timescales. Water Resources Research, 50, 5638–5661. (whereby this study, based on largely similar classes, also had similar conclusions regarding class correlations with e.g. aridity, snowiness, flood timing, low flow timing.) I understand that these studies have already mostly been cited in the main text, but their validity as classifications of seasonal flow regimes that include temporal information has sort of been dismissed by the statement that "The use of catchment characteristics can be problematic because there is often no clear link between these characteristics and streamflow indices". Yet, all of the above-listed studies (except Pardé maybe) show explicitly how their classifications lead to similar within-class behavior of seasonal streamflow regimes. I think there is an opportunity to slightly reframe the paper to acknowledge

that this study complements existing classifications that also incorporate temporal information of flow regimes, (rather than to imply that nothing (useful) exists in this field). (Or alternatively, be more precise and explicit about what the previous classifications can't do that yours does). The use of B-spline basis functions to characterize the streamflow regimes functional behavior in this HESS manuscript seems to be a useful addition to existing literature that I look forward to seeing published in HESS.

Reply: *Dear Wouter, thank you for your thoughts on the framing of our manuscript, which we considered while revising our manuscript. We did by no means intend to imply that nobody has ever looked at timing related streamflow indices in clustering but rather wanted to point out that studies that explicitly consider the temporal information in the continuous signal are rare. This is why we wrote: 'Both the catchment and climate characteristics and the streamflow index approaches neglect nearly all available temporal information embedded in a streamflow time series or regime in the form of temporal (auto-) correlation'. The keywords here are 'nearly' and 'in the form of temporal (auto-) correlation'. We agree that the subsequent sentence may seem exclusive of some contributions and rephrased it to explicitly state that some studies have clustered on streamflow indices related to seasonality and timing: 'While some of the index-based approaches have considered indices related to streamflow timing and seasonality* [*Haines et al.*, 1988; *Bower et al.*, 2004; *McCabe and Wolock*, 2014]*, only very few studies have tried to explicitly take account of temporal streamflow information in clustering hydrological catchments, e.g. by using the shape of the autocorrelation function as an index* [*Toth*, 2013]*, even though such information is potentially very useful.' We also rephrased the sentence on the weak link between certain streamflow and catchment characteristics to: 'The use of catchment characteristics is not always beneficial as certain streamflow indices do not show clear links to these characteristics* [*Ali et al.*, 2012; *Addor et al.*, 2018]*.' We cite the Knoben et al.* [2018] *study under clustering approaches related to climate characteristics as their 'classification scheme is based only on climatic information and can be evaluated with independent streamflow data.' We acknowledge the work by Coopersmith et al.* [2014] *under approaches using streamflow and climate characteristics as they use date of maximum runoff as a streamflow index, which is related to time but does not say anything about the changes of streamflow over time. We acknowledge the work by Berghuijs et al.* [2014] *under the climate indices clustering approaches (formerly mixed approaches). You use a measure for the strength of precipitation seasonality but do not directly include information on the temporal distribution of precipitation over the season. We hope that you find the updated framing more precise and inclusive.*
Modification: p.2: l.53-55, l.47-48

References used in this response to the reviewer

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

Addor, N., G. Nearing, C. Prieto, A. J. Newman, N. Le Vine, and M. P. Clark (2018), A ranking of hydrological signatures based on their predictability in space, *Water Resour. Res.*, *54*, 8792–8812, doi:10.1029/2018WR022606.

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

Berghuijs, W. R., M. Sivapalan, R. A. Woods, and H. H. G. Savenije (2014), Patterns of similiarity of seasonal water balances: A window into streamflow variability over a range of time secales, *Water Resour. Res.*, *50*, 5638–5661, doi:10.1002/2014WR015692.

Bower, D., D. M. Hannah, and G. R. McGregor (2004), Techniques for assessing the climatic sensitivity of river flow regimes, *Hydrol. Process.*, *18*(13), 2515–2543, doi:10.1002/hyp.1479.

Carey, S. K. et al. (2010), Inter-comparison of hydro-climatic regimes across northern catchments: Synchronicity, resistance and resilience, *Hydrol. Process.*, *24*(24), 3591–3602, doi:10.1002/hyp.7880.

Coopersmith, E. J., B. S. Minsker, and M. Sivapalan (2014), Patterns of regional hydroclimatic shifts: An analysis of changing hydrologic regimes, *Water Resoures Researh*, *50*, 1960–1983, doi:10.1111/j.1752-1688.1969.tb04897.x.

Febrero-Bande, M., and M. Oviedo de la Fuente (2012), Statistical Computing in Functional Data Analysis: The R Package fda.usc, *J. Stat. Softw.*, *51*(4), 1-3-, doi:10.18637/jss.v051.i04.

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

Jin, S., L. Yang, P. Danielson, C. Homer, J. Fry, and G. Xian (2013), A comprehensive change detection method for updating the National Land Cover Database to circa 2011, *Remote Sens. Environ.*, *132*, 159–175, doi:10.1016/j.rse.2013.01.012.

Knoben, W. J. M., R. A. Woods, and J. E. Freer (2018), A quantitative hydrological climate classification evaluated with independent streamflow data, *Water Resour. Res.*, *54*(7), 5088–5109, doi:10.1029/2018WR022913.

Lins, H. F. (2012), *USGS Hydro-Climatic Data Network 2009 (HCDN–2009): U.S. Geological Survey Fact Sheet 2012–3047*, Reston, VA.

McCabe, G. J., and D. M. Wolock (2014), Spatial and temporal patterns in conterminous United States streamflow characteristics, *Geophys. Res. Lett.*, *41*(19), 6889–6897, doi:10.1002/2014GL061980.

Melsen, L., N. Addor, N. Mizukami, A. Newman, P. Torfs, M. Clark, R. Uijlenhoet, and R. Teuling (2018), Mapping (dis) agreement in hydrologic projections, *Hydrol. Earth Syst. Sci.*, *22*, 1775–1791, doi:10.5194/hess-22-1775-2018.

Newman, A. J. et al. (2015), Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: Data set characteristics and assessment of regional variability in hydrologic model performance, *Hydrol. Earth Syst. Sci.*, *19*(1), 209–223, doi:10.5194/hess-19-209-2015.

Ramsay, J. O., H. Wickham, S. Graves, and G. Hooker (2014), Package “fda”: Functional data analysis, *CRAN*.

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

Future streamflow regime changes in the United States: assessment using functional classification

Manuela I. Brunner[1], Lieke A. Melsen[2], Andrew J. Newman[1], Andrew W. Wood[3,1], and Martyn P. Clark[4]

[1]National Center for Atmospheric Research (NCAR), Boulder CO, United States
[2]Hydrology and Quantitative Water Management, Wageningen University, Wageningen, Netherlands
[3]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research (NCAR), Boulder CO, United States
[4]College of Arts and Science, University of Saskatchewan, Canmore, Canada

Correspondence: Manuela I. Brunner (manuelab@ucar.edu)

Abstract. Streamflow regimes are changing and expected to further change under the influence of climate change with potential impacts on flow variability and the seasonality of extremes. However, not all types of regimes are going to change in the same way. Climate change impact assessments can therefore benefit from identifying classes of catchments with similar streamflow regimes. Traditional catchment classification approaches have focused on specific meteorological and/or streamflow indices usually neglecting the temporal information stored in the data. The aim of this study is two-fold: (1) develop a catchment classification scheme that enables incorporating such temporal information and (2) use the scheme to evaluate changes in future flow regimes.

We use the developed classification scheme, which relies on a functional data representation, to cluster a large set of catchments in the conterminous United States (CONUS) according to their mean annual hydrographs. We identify five regime classes that summarize the behavior of catchments in the CONUS: 1) *Intermittent regime*, 2) *weak winter regime*, 3) *strong winter regime*, 4) *New Year's regime*, and 5) *melt regime*. Our results show that these spatially contiguous classes are not only similar in terms of their regimes, but also their flood and drought behavior, as well as their physiographical and meteorological characteristics. We therefore deem the functional regime classes valuable for a number of applications going beyond change assessments including model validation studies or predictions of streamflow characteristics in ungauged basins.

To assess future regime changes, we use simulated discharge time series obtained from the Variable Infiltration Capacity hydrologic model driven with meteorological time series generated by five general circulation models. A comparison of the future regime classes derived from these simulations with current classes shows that robust regime changes are expected only for currently melt-influenced regions in the Rocky Mountains. These changes in mountainous, upstream regions may require adapting water management strategies to ensure sufficient water supply in dependent downstream regions.

keypoints:

1. Functional data clustering enables forming clusters of catchments with similar hydrological regimes and a similar drought and flood behavior.

2. We identify five streamflow regime clusters: 1) Intermittent regime, 2) weak winter regime, 3) strong winter regime, 4) New Year's regime, and 5) melt regime.
3. Future regime changes are most pronounced for currently melt-dominated regimes in the Rocky Mountains.
4. Functional regime clusters have widespread utility for predictions in ungauged basins and hydroclimate analyses.

keywords: functional data analysis, clustering, climate change, prediction in ungauged basins, CAMELS, random forest

1 Introduction

The characteristics of streamflow regimes, as here described by mean annual hydrographs, include streamflow variability and seasonality and influence the hydrological functioning of a catchment. Such regimes are undergoing changes and expected to further change under future climate conditions (Addor et al., 2014; Arnell, 1999; Brunner et al., 2019b; Horton et al., 2006; Laghari et al., 2012; Leng et al., 2016; Milano et al., 2015). Regime changes are caused by changes in precipitation seasonality and intensity (Brönnimann et al., 2018) and seasonal shifts and decreases in melt contributions (Stewart et al., 2005; Farinotti et al., 2016; Jenicek et al., 2018) related to reduced snow and glacier storage (Beniston et al., 2018; Mote et al., 2005, 2018). Predicted regime changes are relatively robust (Addor et al., 2014) compared to changes in high and low flows, which are highly uncertain (Brunner et al., 2019c; Madsen et al., 2014) because of diverse uncertainty sources introduced in various steps along the modeling chain (Clark et al., 2016). It has been shown that future regime changes can be linked to changes in flood and drought characteristics, e.g. the seasonality and magnitude of floods (Middelkoop et al., 2001) or the duration of droughts (Brunner and Tallaksen, 2019). Quantifying hydrological regime changes can assist in inferring changes in extremes and is crucial for adapting water management practices (Clarvis et al., 2014).

We can improve our understanding of regime changes by employing regime classification in climate change impact assessments (Coopersmith et al., 2014). Most existing (regime) clustering approaches focus on a set of indices either referring to certain physiographical or climatological catchment characteristics (Berghuijs et al., 2014; Knoben et al., 2018; Wolock et al., 2004), specific streamflow indices (Archfield et al., 2014; Bower et al., 2004; Haines et al., 1988; McCabe and Wolock, 2014), or a mixture of the two (Coopersmith et al., 2012; Kuentz et al., 2017; McManamay and Derolph, 2019; Sawicz et al., 2011; Sharghi et al., 2018; Wagener et al., 2007). The use of catchment characteristics is not always beneficial as certain streamflow indices do not show clear links to these characteristics (Ali et al., 2012; Addor et al., 2018; Oudin et al., 2010). One may therefore prefer to work with streamflow indices directly when identifying catchment classes with a similar streamflow behavior. However, the use of streamflow indices requires the subjective choice of streamflow indices which may not fully capture the catchment behavior. Both the catchment and climate characteristics and the streamflow index approaches neglect nearly all available temporal information embedded in a streamflow time series or regime in the form of temporal (auto-) correlation. While some of the index-based approaches have considered indices related to streamflow timing and seasonality (Bower et al., 2004; Haines et al., 1988; McCabe and Wolock, 2014), only very few studies have tried to explicitly take account of temporal streamflow information in clustering hydrological catchments, e.g. by using the shape of the autocorrelation function as an index (Toth, 2013), even though such information is potentially very useful. We here explore how we can make better use of the seasonal and temporal information stored in the hydrological regime using a functional data representation going beyond considering a set of indices.

In contrast to classical multivariate data, functional data are continuously defined (Ramsay and Silverman, 2002). Functional data analysis represents each hydrological regime as a function and therefore circumvents the choice of individual hydrograph characteristics, which enables exploiting the full information stored in the time series or annual hydrograph when clustering catchments (Chebana et al., 2012; Ternynck et al., 2016). The functional form of the data is derived from discrete observations (Ramsay and Silverman, 2002) either by smoothing the data non-parametrically (Jacques and Preda, 2014) or by projecting the data onto a set of basis functions. The basis function (e.g. B-spline, Fourier, or wavelet bases) coefficients can be used for clustering (Cuevas, 2014). It has been shown in previous studies that functional data representations can be beneficial to identify groups of similar hydrographs over a range of temporal scales, such as spring flood events (duration of six months; Ternynck et al., 2016), flood events (duration of several days; Brunner et al., 2018), low flow events (Laaha et al., 2017), diurnal discharges (duration of one day; Hannah et al., 2000), and yearly hydrographs (Merleau et al., 2007; Jamaludin, 2016).

These previous studies focused on a limited number of stations and on current climate conditions. The goals of this study are therefore two-fold: (1) to develop a catchment classification scheme for streamflow regimes useful in climate change impact assessments; and (2) to use this scheme to evaluate changes in future flow regimes.

We develop the catchment classification scheme for a large dataset of 671 catchments over the United States (Newman et al., 2015; Addor et al., 2017) using a functional representation of mean annual hydrographs. This scheme makes better use of the seasonal and temporal information stored in the hydrological regime than index-based approaches and is solely based on streamflow information (i.e. no climatological information is used). However, it does neither consider streamflow patterns at short, event time scales such as flashiness, nor at time scales longer than a year such as interannual variability.

In order to assess future regime changes, we use streamflow time series simulated with the hydrological Variable Infiltration Capacity (VIC) model driven by meteorological data derived from five general circulation models (GCMs) under a high emission scenario. We compare current and future regime class memberships to identify catchments with future regime changes. Such change assessments are of paramount importance in preparing for future water management strategies because future regime shifts can influence the variability and timing of high and low flows.

2 Data and Methods

2.1 Data

We form regime clusters, i.e. clusters of catchments with similar mean annual hydrographs, using observed streamflow data of 671 catchments in the conterminous United States (CONUS) (Newman et al., 2015). The catchments belong to the HCDN-2009 network (Lins, 2012), which consists of a set of stations deemed suitable for analyzing hydrologic variations and trends in a climatic context as flow conditions are undisturbed by artificial diversions and storage and show less than 5% imperviousness as measured by the National Land Cover Database (Jin et al., 2013). The data were downloaded for the period 1981–2018 from the USGS website https://waterdata.usgs.gov/nwis (R-package dataRetrieval; De Cicco et al., 2018) as data for this period was available for most stations in the dataset. In contrast, the regime change analysis uses streamflow simulated by the hydrological Variable Infiltration Capacity (VIC) model for a subset of 605 catchments, for which reliable data on catchment area was available at the time the simulations were produced (Melsen et al., 2018). Kling-Gupta efficiencies obtained over these basins with VIC varied from a first quartile of 0.47, a median of 0.6 and a third quartile of 0.71, with the lowest values obtained in the Great Plains. Physiographical and meteorological characteristics for these catchments are available via the Catchment Attributes and MEteorology for Large-sample Studies data set (CAMELS) (Addor et al., 2017).

2.2 Regime clustering and classification

Hydrological regime clusters are derived using functional data analysis on the observed hydrological regimes of the 671 catchments (Fig. 1). In the functional data framework, each hydrological regime is considered to be a function (Ramsay and Silverman, 2002). To achieve such a functional data representation, we project the discrete observations, i.e. the mean annual hydrographs at daily resolution, to a set of B-spline basis functions (R-package fda; Ramsay et al., 2014) (see illustration in Figure 1.1 a–c) because B-splines are able to mimic the main characteristics of hydrological regimes (Brunner et al., 2018). A (smoothing) spline function is defined by its order of polynomial segments and its number and placement of knots. The number of knots determines the ability of spline functions to represent sharp features in a curve and the knots can be placed such that they are denser in areas with stronger variations than in smooth areas (Höllig and Hörner, 2013). We here use five spline basis functions of order four, which corresponds to a minimal number of basis functions which still allows for flexibility in representing diverse shapes of regimes. The suitability of five spline basis functions is confirmed by the overall silhouette width, which is for more spline basis functions (6 to 10) lower or very similar to the one for five basis functions. The projection of the observed regimes to the five basis functions results in five coefficients per observed regime, one per spline base. The analysis is performed in R using the packages *fda.usc* (Febrero-Bande and Oviedo de la Fuente, 2012) and *fda* (Ramsay et al., 2014) and the following functions: (1) converting regimes to functional data objects: *fdata*, (2) creating B-spline basis functions: *create.bspline.basis*, (3) computing spline coefficients for all regimes: *Data2fd*.

The clustering into regime classes is performed using the R-package *stats* (R Core Team, 2019). A Euclidean distance matrix is computed using the matrix of $n = 671 \times 5$ spline coefficients (Figure 1.2 a–b) (*dist*). We use a hierarchical clustering algorithm (*hclust*) allowing for non-elliptical clusters (Gordon, 1999) with Ward's minimum variance criterion, which minimizes the total within-cluster variance (Ward, 1963). To identify an optimal number of clusters, we cut the tree at $k = 2, ..., 30$ clusters (*cutree*) and compute the mean silhouette width (Rousseeuw, 1987), which provides a measure of clustering validity, for the different numbers of clusters. We finally determine five regime clusters because the mean silhouette width values stabilize at five clusters. A comparison with regime clusters derived by k-means clustering shows that the final clusters formed are relatively stable independent of the choice of the clustering technique. Each of the clusters can be summarized by its median regime identified using the h-mode depth which allows for ordering the regimes within a cluster (Cuevas et al., 2007).

To assess whether the similarities of the catchments within a cluster go beyond their regime type, we compare their physiographical (latitude, area, elevation), climatological (mean precipitation, fraction of snow, aridity), and flood and streamflow drought characteristics. The flood and drought characteristics are determined using a peak-over-threshold (Lang et al., 1999) and a threshold-level approach (Yevjevich, 1967), respectively. The flood threshold is fixed at the 25th percentile of the annual maxima time series of each catchment separately to guarantee a balanced number of extracted events across catchments (Schlef et al., 2019). The drought threshold is fixed at the highest value of the annual minimum time series and the time series smoothed over a window of 30 years to limit the extraction of dependent events (Brunner et al., 2019d; Tallaksen and Hisdal, 1997).

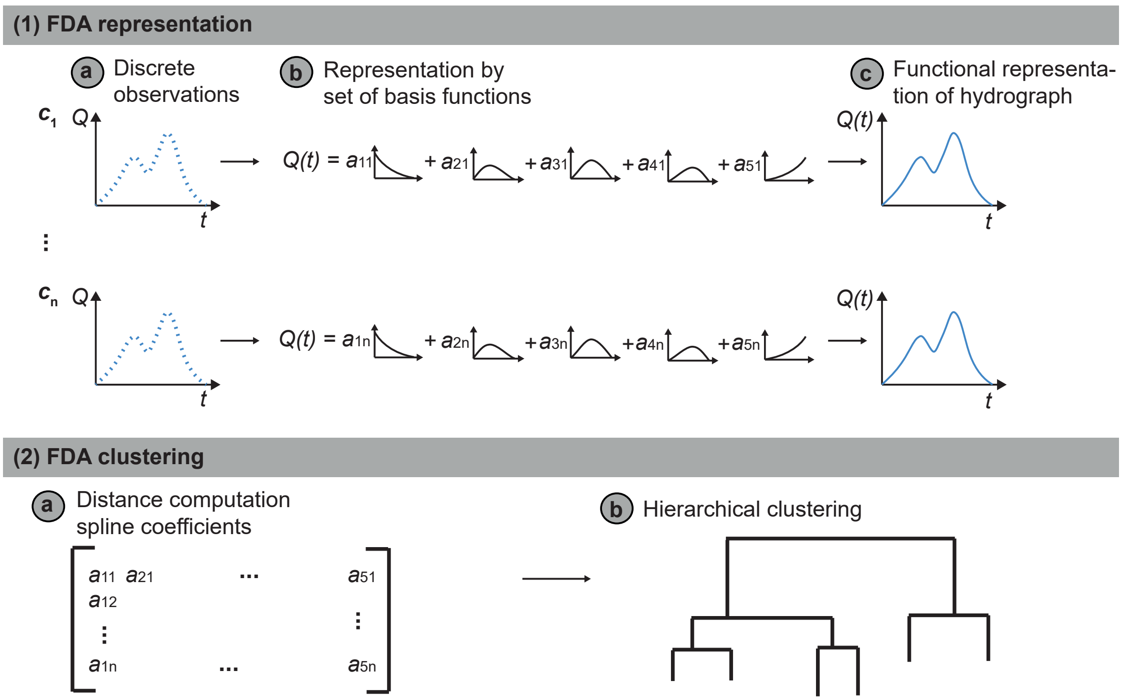

Figure 1. Functional data (FDA) clustering procedure: (1) FDA representation of regimes by projecting (a) discrete observations to (b) a set of spline bases to derive a (c) functional representation of the hydrological regimes. (2) FDA clustering by computing (a) a distance matrix using the spline coefficients from Step 1 in a (b) hierarchical clustering procedure.

To further investigate the physiographical and climatological controls on regime class membership and to check whether regime classes can potentially be predicted for ungauged catchments, we perform a random forest classification (Breiman, 2001; Harrell, 2015; James et al., 2013). We fit the model to 33 non-hydrological catchment characteristics in the CAMELS dataset (Addor et al., 2017), i.e. topographical, soil, geological, and climatological characteristics, excluding gauge ids and characteristics with missing values (‘second most common geologic class in the catchment’ and ‘subsurface porosity’), to predict regime class membership (R-package randomForest; Liaw and Wiener, 2002). The related analysis of estimated variable importance allows for identifying factors important in determining regime class membership, which is useful for ungauged basins where the regime class cannot be determined based on discharge observations.

2.3 Model simulations

For the regime change analysis, we use daily streamflow time series simulated by Melsen et al. (2018) in a model intercomparison project. They ran the Hydrologiska Byråns Vattenbalansavdelning model (HBV; Bergström, 1976), the Variable Infiltration Capacity model (VIC; Liang et al., 1994), and the Sacramento Soil Moisture Accounting model (SAC-SMA) combined with SNOW–17 (Newman et al., 2015) for 605 catchments in the CAMELS dataset for a period representing current (1985–2008) and future climate conditions (2070–2100). Each of these models was run with a large number of parameter sets sampled using Sobol-based Latin hypercube sampling (Bratley and Fox, 1988) and forced with daily observed meteorological variables for the current period (Daymet; Thornton et al., 2012). The performance of each of the sampled parameter sets was evaluated by comparing the model simulations with observed discharge data over a 23-year period (1985–2008) (USGS, 2019) using the Kling–Gupta efficiency metric (Gupta et al., 2009) defined as:

$$E_{KG}(Q) = 1 - \sqrt{(\rho - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2}, \quad (1)$$

where ρ is the correlation between observed and simulated runoff, α is the standard deviation of the simulated runoff divided by the standard deviation of observed runoff, and β is the mean of the simulated runoff, divided by the mean of the observed runoff.

Here we focus on the VIC model and those model runs derived using the parameter set resulting in the best model performance in terms of E_{KG}. E_{KG} values over all stations ranged from a first quartile of 0.47 over a median of 0.60 to a third quartile of 0.71 with the lowest values obtained in the Great Plains.

Melsen et al. (2018) forced the VIC model with daily output from General Circulation Models (GCMs) which was statistically downscaled using the bias-correction and spatial disaggregation (BCSD) method of Wood et al. (2004), for both the current and future period (Department of the Interior, Bureau of Reclamation, Technical Services Center, 2013). They used the output of five different climate models from the Coupled Model Intercomparison Project Phase 5 (CMIP5; Taylor et al., 2012) including CCSM4 (ccsm), CNRM-CM5 (cnrm), INM-CM4 (inmcm), IPSL-CM5A-MR (ipsl), and MPI-ESM-MR (mpi), and the Representative Concentration Pathway 8.5 (RCP8.5; Moss et al., 2010) representing a high emission scenario.

We here use three types of model runs; a control run, where the hydrological model is forced with the observed Daymet meteorology (1985–2008); five reference runs, one per GCM, where the hydrological model is forced with the simulated meteorology for current conditions (1985–2008); and five future runs, where the hydrological model is forced with simulated meteorology for the future period (2070–2100). We refer to the regimes derived from the control run as the control regimes, those regimes derived from the reference simulations as the reference regimes, and those regimes derived from the future runs as future regimes.

2.4 Evaluation of simulated regimes

To determine the suitability of the VIC model for representing regime changes, we extend the model evaluation from the Kling-Gupta efficiency E_{KG} (Eq. 1), which provides an integrative measure of model performance, to a climate sensitivity analysis performed on the control run and a comparison of observed and simulated regime classes performed on the control and reference runs. In the climate sensitivity analysis, we assess whether the hydrological model reacts to changes in mean temperature and precipitation in the same way as observations. In terms of precipitation, this corresponds to checking whether the model captures the resistance of a catchment, i.e. the degree to which runoff is coupled with precipitation (Carey et al., 2010). To do so, we follow a technique presented in Wood et al. (2004) that involves creating many samples of modeled and observed climate and streamflow, and assessing sensitivities from mean behavior of each sample. The multi-year samples help to average out the confounding effects of other influences, such as the initial catchment moisture in individual years. Accordingly, we generate new temperature, precipitation, and streamflow time series by resampling the available hydrological years with replacement ($n = 5000$ times). We compute mean temperature, precipitation, and streamflow for the resampled time series to derive a relationship between mean streamflow and the two meteorological variables. Conducting this experiment for both observed and simulated time series supports analysis of whether the simulated streamflow time series react to changes in mean annual climate in the same way as observed time series.

To assess the ability of the VIC model to simulate the observed regime class, we compare observed to simulated regime classes for the control and reference runs. To assign simulated regimes to one of the five classes, we fit a second classification model using a random forest, which enables classifying a given mean annual hydrograph into one of the five regime classes using its B-spline coefficients. This analysis is different from the first random forest analysis, which was aimed at identifying catchment and climate characteristics determining class membership. We use 10-fold cross validation (Hastie et al., 2008) to evaluate the capability of the classification model to correctly predict observed regime classes. The cross-validation shows that the regime-class prediction error is only 2% and that the model can be used to predict class memberships of simulated regimes accurately. We compare the observed regime classes to the regime classes predicted with the random forest model for the simulated control regimes. This comparison shows that the VIC model is well capable of simulating hydrological regimes with a correct regime prediction in more than 95% of the catchments. The prediction error roughly doubles when using the reference instead of the control regimes indicating that additional uncertainty is introduced by using the GCM simulations as meteorological forcing.

2.5 Future regimes

We use the hydrological model simulations to assess regional changes in regime class memberships. To do so, we predict the regime classes for the five reference regimes (one per GCM) and the corresponding future regimes using the random forest classification model. We then compare the predicted future classes to the class of the corresponding reference simulation using a contingency table of counts. We look at the (dis-) agreement of predicted regime changes for the five GCMs and evaluate whether and where most models agree on regime change.

3 Results

3.1 Hydrological regime clusters

Based on the functional data clustering, the hydrological regimes of the 671 catchments in the CAMELS dataset are divided into five clusters resulting in five mostly spatially contiguous regions of catchments with similar annual hydrographs (Fig. 2).

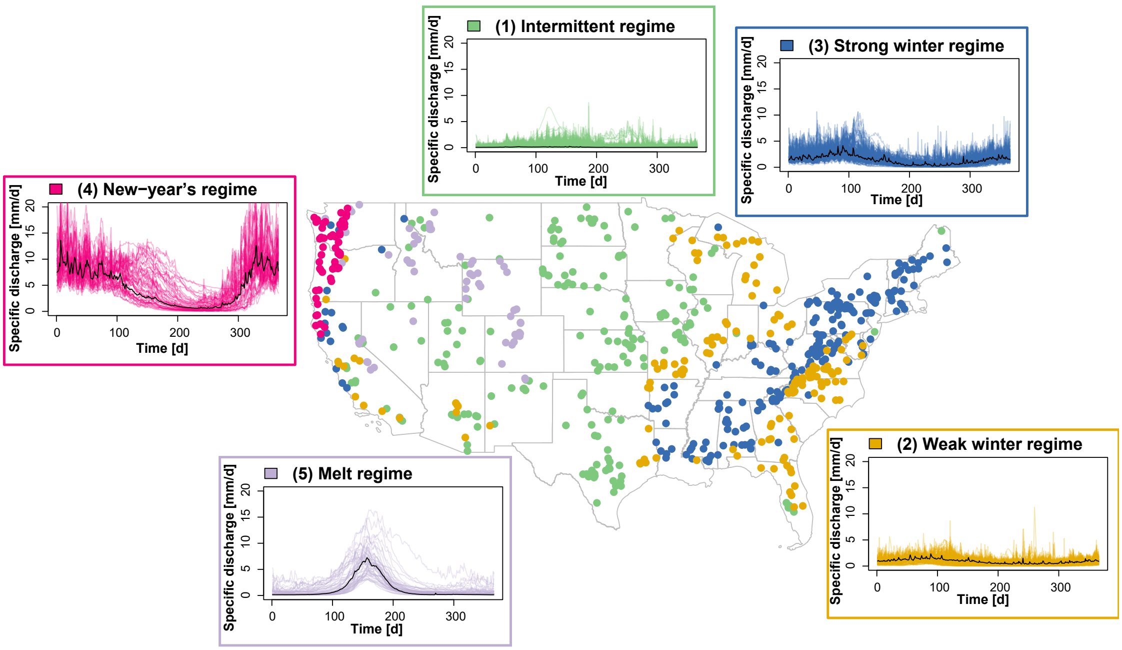

Figure 2. Map of regime clusters and the regimes of the catchments belonging to the five clusters: 1) Intermittent regime, 2) weak winter regime, 3) strong winter regime, 4) New Year's regime, and 5) melt regime. Regimes of individual catchments are colored according to their cluster membership and the median hydrograph per cluster is given in black.

1. The first cluster, which we here call the *intermittent regime* cluster, comprises regimes with a very weak seasonality, dominated by the occurrence of short precipitation events related to thunderstorms or fronts. The catchments belonging to this region mostly lie in the Great Plains, the Great Basin, and the Plateau region (158 catchments).

2. The second cluster, here referred to as *weak winter regime*, comprises regimes showing a weak seasonality with slightly more discharge in winter and spring than in summer and fall. The catchments belonging to this cluster lie in the Coastal Plain, the Lake region, and parts of the Prairie region (127 catchments).

3. The third cluster, i.e. the *strong winter regime*, is similar to the previous regime type with higher winter and spring discharge compared to summer and fall but a slightly more expressed seasonality. The catchments in this cluster mostly belong to the Appalachian region (206 catchments).

4. The catchments in the fourth cluster, which we call *New Year's regime*, have a very strong seasonality with high discharge in winter in general and around New Year in particular, but low discharge in summer. Catchments in this region are located in the Pacific Northwest (57 catchments).

5. The fifth cluster comprises regimes that are snowmelt-dominated and show high discharge in spring and summer vs. low discharge in winter and fall. The catchments belonging to this *melt regime* are located in the Rocky Mountains (57 catchments).

The regime classes are provided for the 671 catchments in the CAMELS dataset via HydroShare (Brunner, 2020): https://doi.org/10.4211/hs.069f552f96ef4e638f4bec281c5016ad.

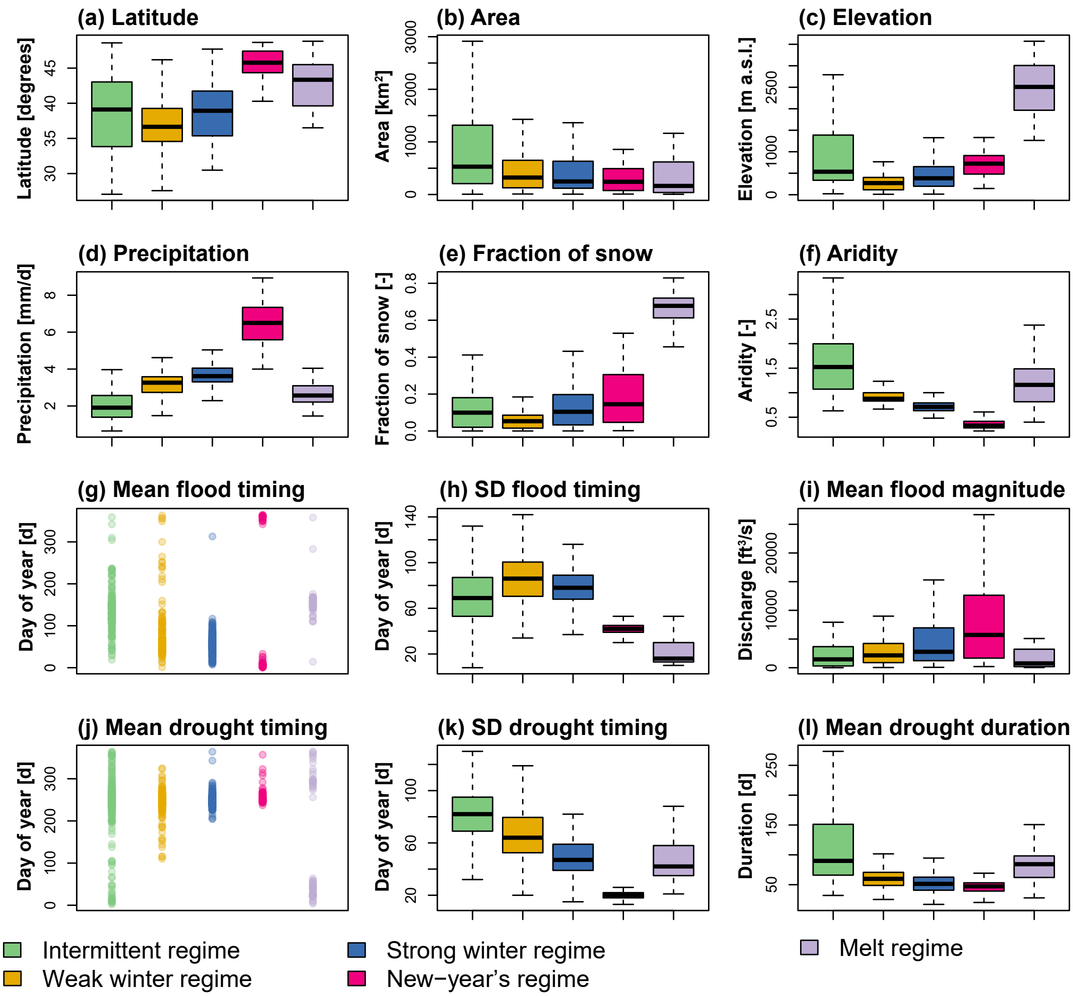

Figure 3. Catchment (a–f), flood (g–i), and drought characteristics (j–l) of the catchments belonging to the five regimes: intermittent, weak winter, strong winter, New-year's, and melt. Characteristics (a–f) were derived from the CAMELS dataset, the flood and drought characteristics using a peak-over-threshold/threshold-level approach, respectively. SD = standard deviation. The black lines in the boxplot indicate the median, the upper and lower whiskers correspond to $1.5 * R_{\mathrm{IQ}}$, where R_{IQ} is to the inter-quartile range. Outliers are not displayed.

The catchments in the five regime clusters are not only similar in terms of their regimes, according to which the clusters were formed, but also in terms of their physiographical, climatological, and flood and drought characteristics (Fig. 3).

Catchments with an *intermittent regime* are comparably large, receive only small precipitation amounts, and are dry. Floods occur mainly in spring and summer while droughts occur in fall and winter. Flood magnitudes are comparably small while droughts are longer than droughts of catchments belonging to other regime clusters. Catchments with a *weak winter regime* lie at low elevations and only a small fraction of total discharge is contributed by snow. These catchments show flood occurrence in winter and spring and droughts in fall. Catchments with a *strong winter regime* lie at relatively low elevations and receive a medium amount of precipitation. Floods occur in winter and droughts in fall. Compared to catchments with a weak winter regime, catchments with a strong winter regime lie at higher elevations, show higher fractions of snow and are characterized by larger flood magnitudes. Catchments with a *New Year's regime* lie at high latitudes and receive a lot of precipitation. Floods occur around New Year and droughts in late fall. Flood magnitudes are very pronounced. Catchments with a *melt-dominated regime* lie at high elevations and a large part of their discharge is melt water. Floods in these catchments occur in spring and early summer due to melt processes and droughts occur in the winter months due to snow accumulation.

The random forest classification model fitted to the regime clusters and a variety of physiographical and climatological catchment characteristics allows for reliable predictions of the correct regime class (prediction error 10%) based on catchment characteristics only. The related variable importance analysis shows that the most important variables for predicting regime classes are climatological characteristics including mean precipitation and aridity. Important physiographical predictors include the longitude and latitude of the gauge location and catchment mean slope and elevation. Excluding these physiographical explanatory variables from the random forest model results in only a small decrease in prediction performance (prediction error 12%).

3.2 Model validation

Before simulations are used to investigate changes in streamflow regimes, we tested whether climate sensitivity is realistically mimicked by the applied model. The simulated time series show a similar reaction of mean discharge to changes in mean temperature and precipitation as the observed series (Fig. 4).

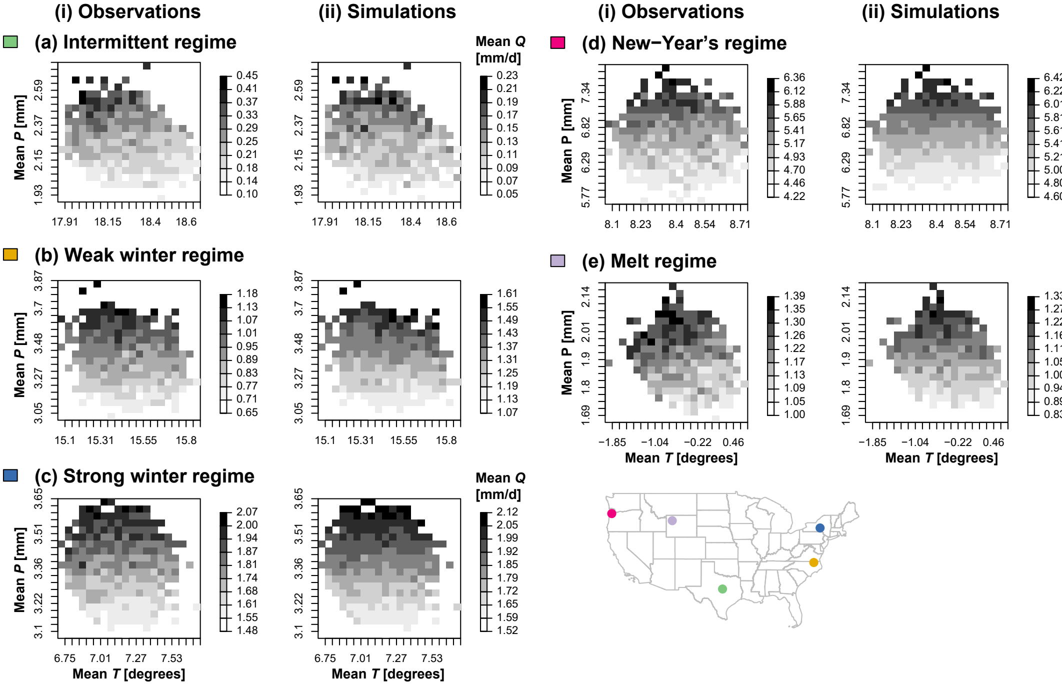

Figure 4. Climate sensitivity analysis for observations (i) and simulations (ii): Dependence of mean discharge (Q) on precipitation (P) and temperature (T) for five example catchments, one per regime class: (a) intermittent regime: Cowhouse creek at Pidcoke, TX (USGS 08101000); (b) weak winter regime: Potecasi creek near Union, NC (USGS, 02053200); (c) strong winter regime: Otselic river at Cincinnatus NY (USGS 01510000); (d) New Year's regime: Tucca creek near Blaine, OR (USGS 14303200); and (e) melt regime: South Fork Shoshone river near Valley, WY (USGS 06280300).

Higher mean precipitation leads to higher mean discharge independent of the catchment and regime. The reaction of streamflow to temperature, however, seems to depend on the catchment because the relationship between mean temperature and mean discharge is generally weak and can be positive or negative. Based on a visual analysis, the realistic simulation of climate sensitivities of mean discharge by the VIC model make it a suitable choice for climate impact assessments of regimes. A quantitative comparison of gradients in these response surfaces over all catchments confirms that the observed and modeled temperature sensitivities are weak while precipitation sensitivities are similar (Kolmogorov–Smirnov test does not reject the null hypothesis that observed and simulated gradients were drawn from the same continuous distribution at level of significance $\alpha = 0.05$). The sensitivity gradients are computed on the response surface of each catchment in the horizontal direction for temperature and in the vertical direction for precipitation.

The VIC model is also able to simulate regimes matching the observed regime classes. The classes of the simulated control regimes predicted using the random forest classification model match the observed regime classes in more than 95% of the catchments (prediction error <5%). The regime class prediction error almost doubles for the reference regimes (prediction error 8–10%) but still allows for the simulation of the correct regime class in more than 90% of the catchments. The good match of simulated control and reference regimes with the observed regimes is illustrated in Fig. 5 for the example catchments with a weak and strong winter regime, a New-Year's, and a melt regime (b–e). In contrast, the regime of the example catchment with an intermittent regime is poorly simulated (a).

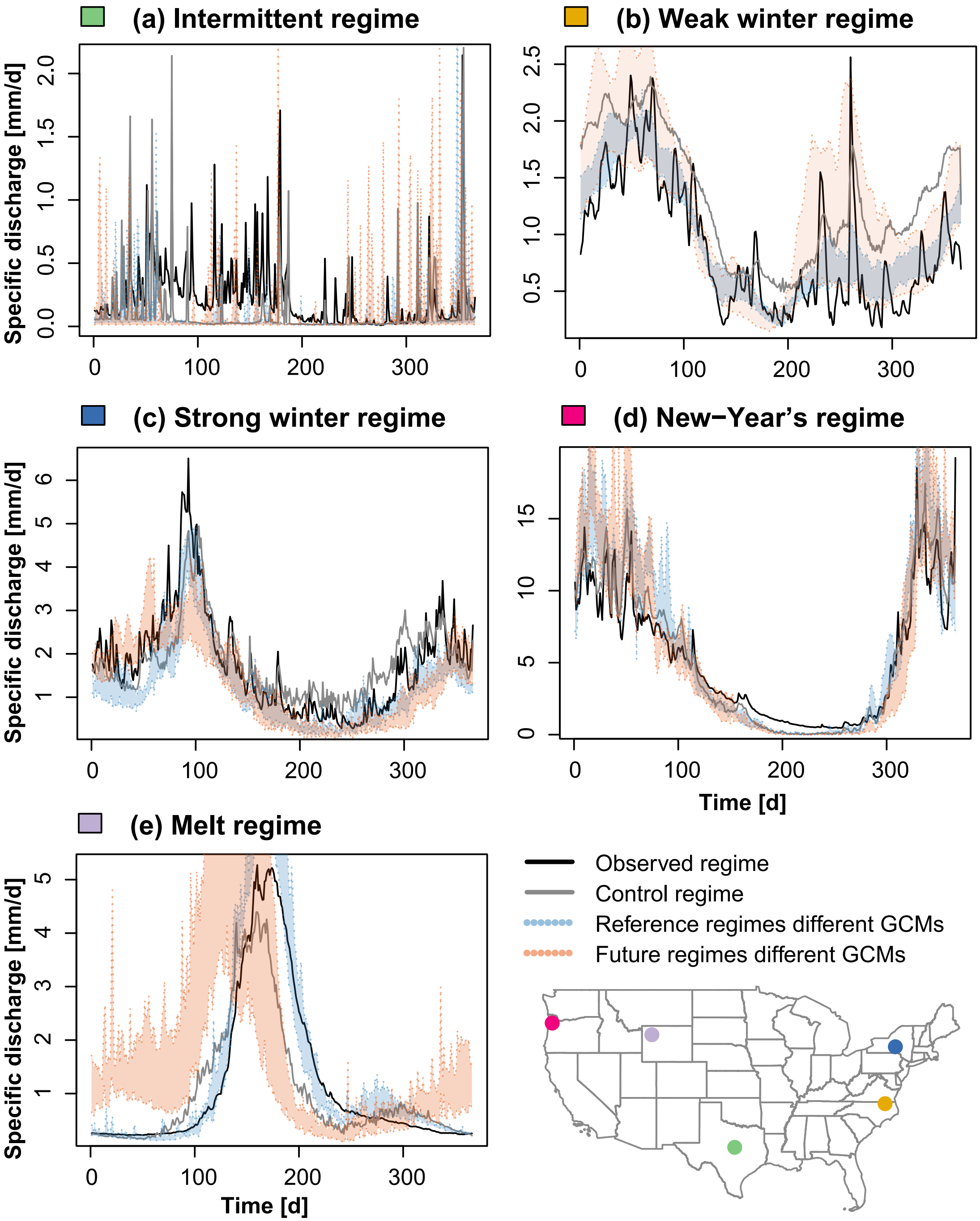

Figure 5. Comparison of observed (black) and simulated control regimes (observed meteorology; grey) with simulated reference (1981-2008; blue) and future regimes (2070-2100; red) derived from the five GCMs for the five example catchments, one per regime type: (a) intermittent regime: Cowhouse creek at Pidcoke, TX (USGS 08101000); (b) weak winter regime: Potecasi creek near Union, NC (USGS, 02053200); (c) strong winter regime: Otselic river at Cincinnatus NY (USGS 01510000); (d) New Year's regime: Tucca creek near Blaine, OR (USGS 14303200); and (e) melt regime: South Fork Shoshone river near Valley, WY (USGS 06280300).

The results of our model evaluation show that the VIC model performs well in simulating the correct regime types when forced with observed meteorological data and in simulating changes in mean discharge as a response to changes in mean temperature and precipitation. However, simulating the observed regime classes becomes more difficult when forcing the model with simulated meteorological data generated by GCMs, in particular in certain areas in the Midwest, in the Pacific Northwest, and a few catchments in the Rocky Mountains and Florida. Over all catchments, regimes of catchments with a weak winter regime and an intermittent regime, i.e. regimes with a weak seasonality are not well reproduced in GCM-forced simulations (Fig. 6 left bars). In contrast, regimes with a strong seasonality such as strong winter, New Year's, and melt regimes are well simulated. These results highlight that model performance depends on regime type.

3.3 Future regimes

Our results show that streamflow regimes may be subject to future changes. This is illustrated by the regime shift of the catchment with a melt regime in Figure 5e. However, these regime shifts do not affect all catchments and are to some extent dependent on the GCM and regime considered (Fig. 6). Only few regime changes are expected for catchments with a currently intermittent, strong winter, and New Year's regime. Moderate regime changes are predicted for catchments with a currently weak winter regime, however, simulation error is quite large for this type of regime. The biggest changes are predicted for currently melt-dominated regimes while catchments with current New Year's regimes hardly change. Currently intermittent regimes are mostly changing to weak winter regimes, currently weak winter regimes to intermittent or strong winter regimes, and currently strong winter regimes to weak winter or New Year's regimes, regime types relatively close to their current regime. In contrast, melt regimes can change into any type of regime depending on the local climate. Catchments without predicted regime changes may still undergo changes in individual streamflow characteristics such as variability or low and high flows.

Geographically, regime changes are expected according to most GCMs in the Rocky and Appalachian Mountains and to a lesser degree in the Pacific Northwest and the Midwest. In contrast, regimes of catchments in the Great Plains are predicted to be mostly unaffected by changes. These results are summarized in Figure 7a where all catchments with at least one GCM predicting future regime changes are colored according to their current regime type. Even if all GCMs agree on changes, they may not agree on the direction of change (Fig. 7b). Catchments where models agree both on changes and their direction are mostly located in the Rocky Mountains. The currently melt-dominated regimes are expected to change to regimes with less discharge in summer and more discharge in winter. In all other regions, at least one model deviates from the majority regime prediction and the direction of change is less clear.

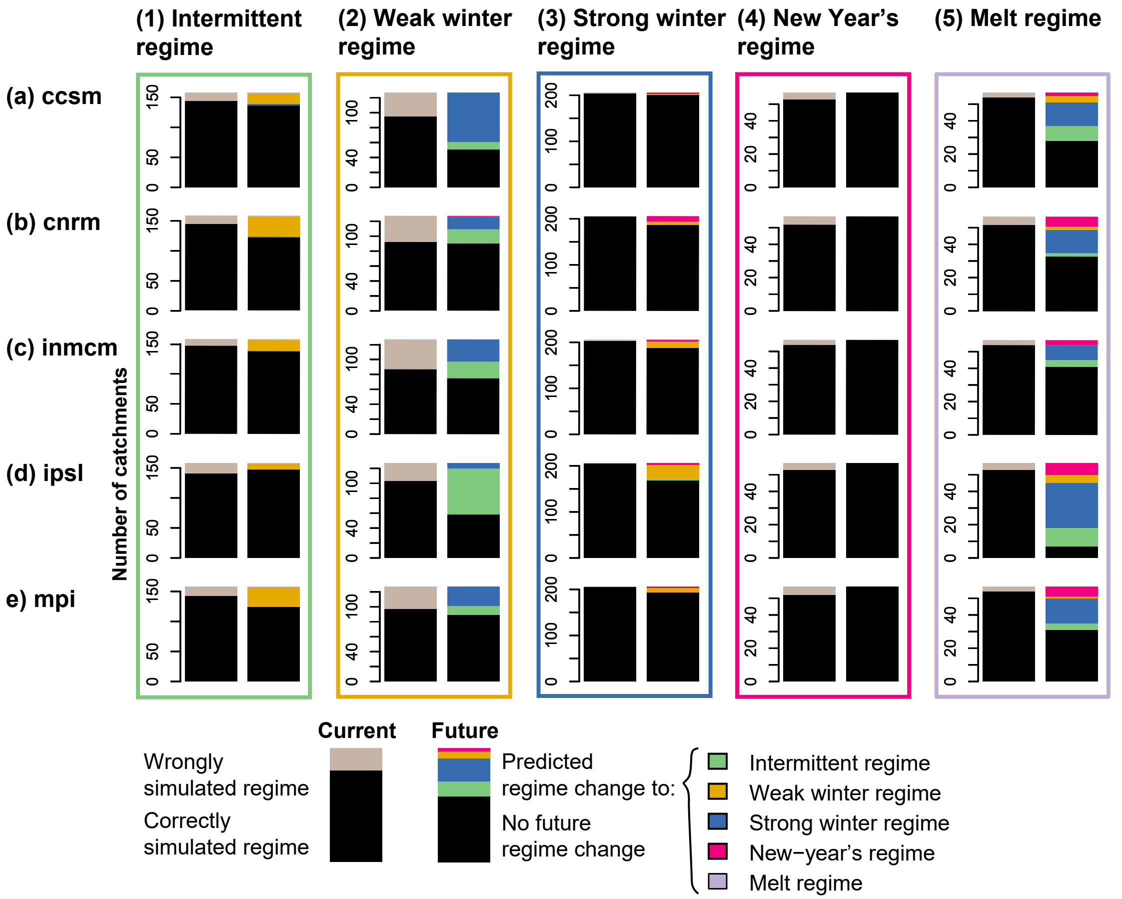

Figure 6. Current regime simulation error and future predicted regime changes for the five regimes: (1) Intermittent, (2) weak winter regime, (3) strong winter regime, (4) New Year's regime, and (5) melt regime and the five GCMs: (a) ccsm, (b) cnrm, (c) inmcm, (d) ipsl, and (e) mpi. The number of catchments where the reference simulations result in the observed regime class and a wrong regime class are given in black and grey, respectively. The number of catchments with no predicted regime changes is given in black, the direction of change for the catchments with predicted changes is indicated by the respective regime color.

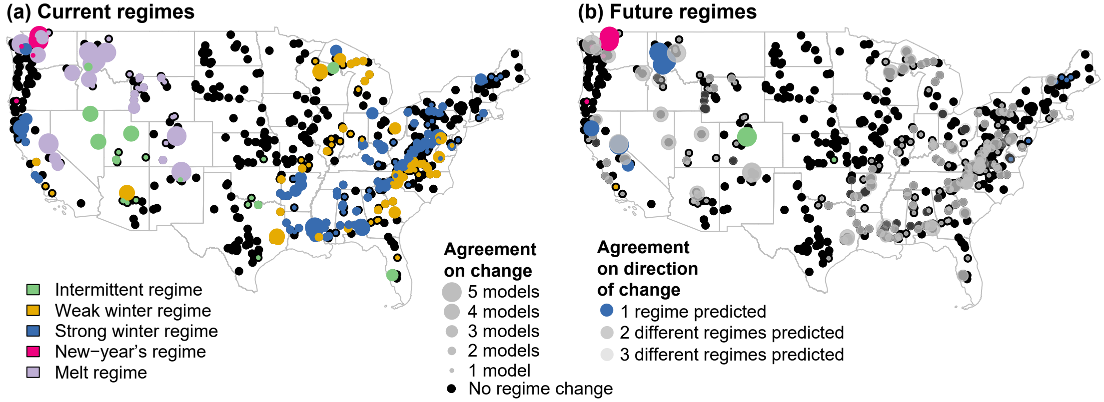

Figure 7. (a) Current regimes and agreement of models regarding regime changes. Catchments colored according to their observed regime show catchments where at least one out of the five GCMs predicts a regime class change. The size of the dot indicates the strength of model agreement. (b) Future regimes and agreement of models regarding the direction of change. The size of the dot indicates agreement on change, the color of the dot the agreement on direction of change. All GCMs predict the same change in colored catchments, GCMs disagree on the direction of change in grey catchments where the shading indicates the strength of agreement. Black catchments are either predicted to experience no changes or their reference regime was incorrectly predicted by more than two GCMs.

4 Discussion

4.1 Hydrological regime clusters

We find functional data clustering to be a useful tool for identifying clusters of catchments with not only similar streamflow regimes but also similar catchment, meteorological, flood and drought characteristics. This similarity corroborates findings by Bower et al. (2004) and McCabe and Wolock (2014) who established a clear link between similarity in streamflow seasonality and climatic and physical similarity. However, it is in contrast to findings by Ali et al. (2012) who found that catchments similar with respect to a set of flow indices are not necessarily physically similar. Explicitly including seasonality or information on the temporal autocorrelation of regimes may therefore help to identify clusters of catchments which are not only hydrologically but also physically similar.

The five regime clusters are mostly spatially contiguous and show similarities to the four catchment clusters built by McManamay and Derolph (2019) who used 110 different hydrological characteristics in their clustering procedure. Our approach circumvents computing and selecting (a large number of) streamflow characteristics by applying the clustering procedure on a functional representation of the mean annual hydrographs directly. The five regime clusters identified also show spatial similarities with the ten catchment clusters formed by Jehn et al. (2020) for the same set of catchments using a small set of hydrological streamflow characteristics. However, our clustering scheme results in larger clusters than the ones seen in Jehn et al. (2020). Similarly to Jehn et al. (2020) and Yaeger et al. (2012), we find that meteorological characteristics in general and mean precipitation and aridity in particular are stronger predictors for hydrological class membership than physiographical catchment characteristics. However, we also find that catchment mean slope, elevation, and location help to explain regime class membership. The relationship of class membership to physiographical characteristics may be weaker than the one to climatic characteristics as the clusters are formed using the mean annual hydrographs whose seasonality is strongly influenced by climate. The link to physiographical characteristics may be stronger if streamflow characteristics at an event time scale are considered.

The strong link between regime classes and meteorological and physiographical catchment characteristics enables attributing ungauged catchments, where streamflow data are not available, to one of the regime classes. This attribution can be achieved by using the first random forest model fitted in this analysis enabling predictions of regime class membership using physiographical and climatological characteristics. The ability to attribute an ungauged catchment to one of the regime classes is potentially very useful to predict streamflow characteristics in ungauged basins.

The streamflow regime classes identified here do not comprise classes of catchments with major flow alterations as the clustering was performed using streamflow regimes from catchments with minimal human impact. The five classes proposed here are therefore of limited use if a problem requires including catchments with strong human flow alterations. A flow regime of a regulated stream may still be attributed to one of the five regime classes identified if the altered regime shows similarities with the flow seasonality and variability of one of the 'natural' classes. However, if flow alteration leads to the emergence of regimes clearly distinct from those observed under natural conditions, additional regime classes would be necessary. In addition, the relationships between catchment characteristics and class memberships would need to be revised to enable the assignment of ungauged catchments to one of the classes in the updated set.

4.2 Model validation

The uncertainty introduced into the simulations by using the GCM meteorology as shown by differences between the downscaled and observed time series could have different reasons. One potential reason for these differences is that the observations used to fit the downscaling model are fairly short. Another reason could be that the downscaling model was fitted using a different dataset (Maurer et al., 2002) than used to calibrate the hydrological model (Thornton et al., 2012) highlighting that precipitation observations are subject to measurement errors.

4.3 Future changes

The future regime changes detected are relatively robust for currently melt-influenced regimes while they are not consistent for the other regime types. The predicted changes in melt influenced regimes are in line with findings by Coopersmith et al. (2014) who found that snow pack has diminished in the Rocky Mountains in the past and are consistent with future predicted increases in temperature (Vose et al., 2017) and related decreases in snowpack (Easterling et al., 2017). In contrast, predicted changes in precipitation are variable in space and time (Easterling et al., 2017), which disables clear change assessments for rainfall-dominated regimes. Similarly, Milner et al. (2017) and Adam et al. (2009) found on a global scale that warming is generally associated with reductions in glaciermelt and losses of snowpack, respectively, and therefore changes in streamflow seasonality.

However, Adam et al. (2009) also point out that catchments more sensitive to changes in precipitation than temperature may show different change patterns. While our study focused on detecting changes between existing regime classes, there might emerge new regimes (Leng et al., 2016), which we have not considered here.

The changes of melt-influenced towards more rainfall-influenced regimes in the Rocky Mountains and the dependence of flood and drought timing on the streamflow regime allows us to think about the impacts of regime changes on future extremes. A shift from a melt regime to one of the rainfall-influenced regimes implies a shift of the flood and drought seasons. Under a melt regime, floods mainly occur in spring and early summer when snowmelt and rain-snow interactions enhance the flood signal. In contrast, droughts are mainly observed in winter due to snow accumulation temporarily storing water in the catchment. A decreased influence of snow therefore moves the flood season away from spring/early summer into the season with the biggest precipitation input, which is often winter or spring. Analogously, the drought season moves away from winter into summer and fall, the seasons with the largest precipitation deficits. At the same time, drought and flood magnitudes may also be impacted, however, the direction of change is less clear there. These expected changes in flood and drought timing and magnitude have potential implications for the predictability of extremes and the spatial coherence in flood and drought occurrence.

5 Conclusions

The aim of this study was to (1) develop a flow regime classification scheme beneficial for climate impact assessments and to (2) use this scheme to evaluate future regime changes. We find that the functional clustering approach applied to classify flow regimes is efficient because it uses the temporal information stored in hydrographs thereby sidestepping the need to compute a (large) set of streamflow indices and enables identifying contiguous regions with similar streamflow regimes. We conclude that the regime behaviour of the 671 US catchments analyzed here can be summarized by five streamflow regime classes: intermittent regime, weak winter regime, strong winter regime, New Year's regime, and melt regime. These classes are not only similar in their regimes but also their physiographical and meteorological characteristics as well as their extreme streamflow behaviour including the timing and magnitude of droughts and floods. Because of these similarities, we deem the regime classes developed in this study beneficial not only for climate impact assessments but also for model validation and development, the improvement of predictions in ungauged basins, and estimating hydrological model parameters.

Our change impact assessment shows that predicted regime changes are robust in only very few catchments due to model disagreement regarding change and its direction. These GCM-introduced uncertainties demonstrate that predicted regime shifts should be evaluated carefully. Independent of the climate model, however, there is a relatively robust change signal for currently melt-influenced regimes in mountainous catchments even though models do not necessarily agree on the direction of change. Such mountainous catchments take an important role as water towers providing essential freshwater resources to downstream regions (Immerzeel et al., 2020; Viviroli et al., 2007). Expected changes in these mountainous regions, which are crucial for water supply, point out the potential need for adaptations of water management strategies. Water may need to be stored in reservoirs during winter in order to sustain current summer flows in dependent downstream catchments (Brunner et al., 2019a).

A careful evaluation of future regime shifts and their uncertainty can guide decision making on water management and attempt to mitigate the negative impacts of climate change.

Data availability. The regime classes derived for the 671 catchments in the CAMELS data set are provided via HydroShare (Brunner, 2020): https://doi.org/10.4211/hs.069f552f96ef4e638f4bec281c5016ad.

Author contributions. MIB developed the concept of the study together with MPC. LAM provided the streamflow simulations and model evaluation statistics. MIB established the regime clusters, performed the climate impact assessment on regimes, and wrote the first draft of the manuscript. AJN and AWW provided interpretations of model performance, and AWW contributed the model sensitivity evaluation concept. All co-authors have revised and edited the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The daily discharge time series used in this study are available via the USGS website: https://waterdata.usgs.gov/nwis and the CAMELS catchment attributes can be downloaded via https://ral.ucar.edu/solutions/products/camels. This work was supported by the Swiss National Science Foundation via a PostDoc.Mobility grant (Number: P400P2_183844, granted to MIB), and the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977. AJN and AWW were also partially supported by the US Army Corps of Engineers Climate Preparedness and Resilience Program. We thank Ulrich Jehn, Genevieve Ali, Wouter Berghuijs, and an anonymous reviewer for their comments.

References

Adam, J. C., Hamlet, A. F., and Lettenmaier, D. P.: Implications of global climate change for snowmelt hydrology in the twenty-first century, Hydrological Processes, 23, 962–972, https://doi.org/10.1002/hyp, 2009.

Addor, N., Rössler, O., Köplin, N., Huss, M., Weingartner, R., and Seibert, J.: Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments, Water Resources Research, 50, 1–22, https://doi.org/10.1002/2014WR015549, 2014.

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, 5293–5313, https://doi.org/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, 54, 8792–8812, https://doi.org/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, https://doi.org/10.1016/j.advwatres.2012.01.008, 2012.

Archfield, S. A., Kennen, J. G., Carlisle, D. M., and Wolock, D. M.: An objective and parsimonious approach for classifying natural flow regimes at a continental scale, River Research and Applications, 30, 1166–1183, https://doi.org/10.1002/rra, 2014.

Arnell, N. W.: The effect of climate change on hydrological regimes in Europe, Global Environmental Change, 9, 5–23, https://doi.org/10.1016/S0959-3780(98)00015-6, 1999.

Beniston, M., Farinotti, D., Stoffel, M., Andreassen, L. M., Coppola, E., Eckert, N., Fantini, A., Giacona, F., Hauck, C., Huss, M., Huwald, H., Lehning, M., López-Moreno, J. I., Magnusson, J., Marty, C., Morán-Tejéda, E., Morin, S., Naaim, M., Provenzale, A., Rabatel, A., Six, D., Stötter, J., Strasser, U., Terzago, S., and Vincent, C.: The European mountain cryosphere: A review of its current state, trends, and future challenges, The Cryosphere, 12, 759–794, https://doi.org/10.5194/tc-12-759-2018, 2018.

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

Bergström, S.: Development and application of a conceptual runoff model for Scandinavian catchments. Swedish Meteorological and Hydrological Institute (SMHI) RHO 7, Tech. Rep. January 1976, Sveriges Meteorologiska och Hydrologiska Institut, Norrköping, 1976.

Bower, D., Hannah, D. M., and McGregor, G. R.: Techniques for assessing the climatic sensitivity of river flow regimes, Hydrological Processes, 18, 2515–2543, https://doi.org/10.1002/hyp.1479, 2004.

Bratley, P. and Fox, B. L.: Algorithm 659: Implementing Sobol's Quasirandom Sequence Generator, ACM Transactions on Mathematical Software (TOMS), 14, 88–100, https://doi.org/10.1145/42288.214372, 1988.

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Brönnimann, S., Rajczak, J., Fischer, E., Raible, C., Rohrer, M., and Schär, C.: Changing seasonality of moderate and extreme precipitation events in the Alps, Natural Hazards and Earth System Sciences, 18, 2047–2056, https://doi.org/10.5194/nhess-18-2047-2018, 2018.

Brunner, M. I.: Streamflow regimes for CAMELS dataset, https://doi.org/10.4211/hs.069f552f96ef4e638f4bec281c5016ad, http://www.hydroshare.org/resource/069f552f96ef4e638f4bec281c5016ad, 2020.

Brunner, M. I. and Tallaksen, L. M.: Proneness of European catchments to multiyear streamflow droughts, Water Resources Research, 55, 8881– 8894, https://doi.org/10.1029/2019WR025903, 2019.

Brunner, M. I., Viviroli, D., Furrer, R., Seibert, J., and Favre, A.-C.: Identification of flood reactivity regions via the functional clustering of hydrographs, Water Resources Research, 54, 2017WR021 650, https://doi.org/10.1002/2017WR021650, 2018.

Brunner, M. I., Björnsen Gurung, A., Zappa, M., Zekollari, H., Farinotti, D., and Stähli, M.: Present and future water scarcity in Switzerland: Potential for alleviation through reservoirs and lakes, Science of the Total Environment, 666, 1033–1047, https://doi.org/10.1016/j.scitotenv.2019.02.169, 2019a.

Brunner, M. I., Farinotti, D., Zekollari, H., Huss, M., and Zappa, M.: Future shifts in extreme flow regimes in Alpine regions, Hydrological Earth System Sciences, 23, 4471–4489, https://doi.org/10.5194/hess-23-4471-2019, 2019b.

Brunner, M. I., Hingray, B., Zappa, M., and Favre, A. C.: Future trends in the interdependence between flood peaks and volumes: Hydro-climatological drivers and uncertainty, Water Resources Research, 55, 1–15, https://doi.org/10.1029/2019WR024701, 2019c.

Brunner, M. I., Liechti, K., and Zappa, M.: Extremeness of recent drought events in Switzerland: Dependence on variable and return period choice, Natural Hazards and Earth System Sciences, 19, 2311–2323, https://doi.org/10.5194/nhess-19-2311-2019, 2019d.

Carey, S. K., Tetzlaff, D., Seibert, J., Soulsby, C., Buttle, J., Laudon, H., McDonnell, J., McGuire, K., Caissie, D., Shanley, J., Kennedy, M., Devito, K., and Pomeroy, J. W.: Inter-comparison of hydro-climatic regimes across northern catchments: Synchronicity, resistance and resilience, Hydrological Processes, 24, 3591–3602, https://doi.org/10.1002/hyp.7880, 2010.

Chebana, F., Dabo-Niang, S., and Ouarda, T. B. M. J.: Exploratory functional flood frequency analysis and outlier detection, Water Resources Research, 48, W04 514, https://doi.org/10.1029/2011WR011040, 2012.

Clark, M. P., Wilby, R. L., Gutmann, E. D., Vano, J. A., Gangopadhyay, S., Wood, A. W., Fowler, H. J., Prudhomme, C., Arnold, J. R., and Brekke, L. D.: Characterizing uncertainty of the hydrologic impacts of climate change, Current Climate Change Reports, 2, 55–64, https://doi.org/10.1007/s40641-016-0034-x, 2016.

Clarvis, M. H., Fatichi, S., Allan, A., Fuhrer, J., Stoffel, M., Romerio, F., Gaudard, L., Burlando, P., Beniston, M., Xoplaki, E., and Toreti, A.: Governing and managing water resources under changing hydro-climatic contexts: The case of the upper Rhone basin, Environmental Science and Policy, 43, 56–67, https://doi.org/10.1016/j.envsci.2013.11.005, 2014.

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, 4467–4482, https://doi.org/10.5194/hess-16-4467-2012, 2012.

Coopersmith, E. J., Minsker, B. S., and Sivapalan, M.: Patterns of regional hydroclimatic shifts: An analysis of changing hydrologic regimes, Water Resoures Researh, 50, 1960–1983, https://doi.org/10.1111/j.1752-1688.1969.tb04897.x, 2014.

Cuevas, A.: A partial overview of the theory of statistics with functional data, Journal of Statistical Planning and Inference, 147, 1–23, https://doi.org/10.1016/j.jspi.2013.04.002, 2014.

Cuevas, A., Febrero, M., and Fraiman, R.: Robust estimation and classification for functional data via projection-based depth notions, Computational Statistics, 22, 481–496, https://doi.org/10.1007/s00180-007-0053-0, 2007.

De Cicco, L. A., Lorenz, D., Hirsch, R. M., and Watkins, W.: dataRetrieval: R packages for discovering and retrieving water data available from U.S. federal hydrologic web services, https://doi.org/10.5066/P9X4L3GE, https://code.usgs.gov/water/dataRetrieval, 2018.

Department of the Interior, Bureau of Reclamation, Technical Services Center: Downscaled CMIP3 and CMIP5 climate projections: Release of downscaled CMIP5 climate projections, comparison with preceding information, and summary of user needs, Tech. rep., Bureau of Reclamation, Denver, Colorado, https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/techmemo/downscaled_climate.pdf, 2013.

Easterling, D., Kunkel, K., Arnold, J., Knutson, T., LeGrande, A., Leung, L., Vose, R., Waliser, D., and Wehner, M.: Precipitation change in the United States., in: Climate Science Special Report: Fourth National Climate Assessment, Volume I, edited by Wuebbles, D., Fahey, D., Hibbard, K., Dokken, D., Stewart, B., and Maycock, T., chap. 7, pp. 207–230, U.S Gobal Change Research Program, Washington, DC, https://doi.org/10.7930/J0H993CC.U.S., 2017.

Farinotti, D., Pistocchi, A., and Huss, M.: From dwindling ice to headwater lakes: could dams replace glaciers in the European Alps?, Environmental Research Letters, 11, 054 022, https://doi.org/10.1088/1748-9326/11/5/054022, 2016.

Febrero-Bande, M. and Oviedo de la Fuente, M.: Statistical Computing in Functional Data Analysis: The R Package fda.usc, Journal of Statistical Software, 51, 1–3, https://doi.org/10.18637/jss.v051.i04, http://www.jstatsoft.org/v51/i04, 2012.

Gordon, A.: Classification, Chapman & Hall/CRC, Boca Raton, 2nd edn., 1999.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

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

Hannah, D. M., Smith, B. P. G., Grunell, A. M., and McGregor, G. R.: An approach to hydrograph classification, Hydrological Processes, 14, 317–338, 2000.

Harrell, F. E.: Regression Modeling Strategies. With applications to linear models, logistic and ordinal regression, and survival analysis, Springer, Cham, 2015.

Hastie, T., Tibshirani, R., and Friedman, J.: The elements of statistical learning, Springer series in statistics, Springer, Stanford, California, 2008.

Höllig, K. and Hörner, J.: Approximation and modeling with B-splines, Society for industrial and applied mathematics, Philadelphia, 2013.

Horton, P., Schaefli, B., Mezghani, A., Hingray, B., and Musy, A.: Assessment of climate-change impacts on alpine discharge regimes with climate model uncertainty, Hydrological Processes, 20, 2091–2109, https://doi.org/10.1002/hyp.6197, 2006.

Immerzeel, W., Lutz, A., Andrade, M., Bahl, A., Biemans, H., Bolch, T., Hyde, S., Brumby, S., Davies, B., Elmore, A., Emmer, A., Feng, M., Fernández, A., Haritashya, U., Kargel, J., Koppes, M., Kraaijenbrink, P., Kulkarni, A., Mayewski, P., Nepal, S., Pacheco, P., Painter, T., Pellicciotti, F., Rajaram, H., Rupper, S., Sinisalo, A., Shrestha, A., Viviroli, D., Wada, Y., Xiao, C., Yao, T., and Baillie, J.: Importance and vulnerability of the world's water towers, Nature, 577, 364–369, https://doi.org/10.1038/s41586-019-1822-y, 2020.

Jacques, J. and Preda, C.: Model-based clustering for multivariate functional data, Computational Statistics & Data Analysis, 71, 92–106, https://doi.org/10.1016/j.csda.2012.12.004, 2014.

Jamaludin, S.: Streamflow profile classification using functional data analysis : A case study on the Kelantan river basin, in: The 3rd ISM international statistical conference, vol. 1842, pp. 1–11, https://doi.org/10.1063/1.4982836, 2016.

James, G., Witten, D., Hastie, T., and Tibshirani, R.: An introduction to statistical learning. With applications in R., Springer, New York, https://doi.org/10.1007/978-1-4614-7138-7, 2013.

Jehn, F. U., Bestian, K., Breuer, L., Kraft, P., and Houska, T.: Using hydrological and climatic catchment clusters to explore drivers of catchment behavior, Hydrology and Earth System Sciences, 24, 1081–1100, https://doi.org/10.5194/hess-24-1081-2020, 2020.

Jenicek, M., Seibert, J., and Staudinger, M.: Modeling of future changes in seasonal snowpack and impacts on summer low flows in Alpine catchments, Water Resources Research, 54, 538–556, https://doi.org/10.1002/2017WR021648, 2018.

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., and Xian, G.: A comprehensive change detection method for updating the National Land Cover Database to circa 2011, Remote Sensing of Environment, 132, 159–175, https://doi.org/10.1016/j.rse.2013.01.012, 2013.

Knoben, W. J., Woods, R. A., and Freer, J. E.: A quantitative hydrological climate classification evaluated with independent streamflow data, Water Resources Research, 54, 5088–5109, https://doi.org/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, 2863–2879, https://doi.org/10.5194/hess-21-2863-2017, 2017.

Laaha, G., Gauster, T., Tallaksen, L. M., Vidal, J.-P., Stahl, K., Prudhomme, C., Heudorfer, B., Vlnas, R., Ionita, M., Lanen, H. A. J., Adler, M.-J., Caillouet, L., Delus, C., Fendekova, M., Gailliez, S., Hannaford, J., Kingston, D., Loon, A. F. V., Mediero, L., Osuch, M., Romanowicz, R., Sauquet, E., Stagge, J. H., and Wong, W. K.: The European 2015 drought from a hydrological perspective, Hydrology and Earth System Sciences, 21, 3001–3024, https://doi.org/10.5194/hess-21-3001-2017, 2017.

Laghari, A. N., Vanham, D., and Rauch, W.: To what extent does climate change result in a shift in Alpine hydrology? A case study in the Austrian Alps, Hydrological Sciences Journal, 57, 103–117, https://doi.org/10.1080/02626667.2011.637040, 2012.

Lang, M., Ouarda, T., and Bobée, B.: Towards operational guidelines for over-threshold modeling, Journal of Hydrology, 225, 103–117, 1999.

Leng, G., Huang, M., Voisin, N., Zhang, X., Asrar, G. R., and Leung, L. R.: Emergence of new hydrologic regimes of surface water resources in the conterminous United States under future warming, Environmental Research Letters, 11, 1–13, https://doi.org/10.1088/1748-9326/11/11/114003, 2016.

Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, Journal of Geophysical Research, 99, 14 415, https://doi.org/10.1029/94JD00483, 1994.

Liaw, A. and Wiener, M.: Classification and regression by randomForest, R News, 2, 18–22, http://cran.r-project.org/doc/Rnews/, 2002.

Lins, H.: USGS Hydro-Climatic Data Network 2009 (HCDN–2009): U.S. Geological Survey Fact Sheet 2012–3047, Tech. Rep. April, U.S. Geological Survey, Reston, VA, http://pubs.usgs.gov/fs/2012/3047/, 2012.

Madsen, H., Lawrence, D., Lang, M., Martinkova, M., and Kjeldsen, T.: Review of trend analysis and climate change projections of extreme precipitation and floods in Europe, Journal of Hydrology, 519, 3634–3650, https://doi.org/10.1016/j.jhydrol.2014.11.003, 2014.

Maurer, E. P., Wood, A. W., Adam, J. C., Lettenmaier, D. P., and Nijssen, B.: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States, Journal of Climate, 15, 3237–3251, 2002.

McCabe, G. J. and Wolock, D. M.: Spatial and temporal patterns in conterminous United States streamflow characteristics, Geophysical Research Letters, 41, 6889–6897, https://doi.org/10.1002/2014GL061980, 2014.

McManamay, R. A. and Derolph, C. R.: Data descriptor: A stream classification system for the conterminous United States, Scientific Data, 6, 1–18, https://doi.org/10.1038/sdata.2019.17, 2019.

Melsen, L., Addor, N., Mizukami, N., Newman, A., Torfs, P., Clark, M., Uijlenhoet, R., and Teuling, R.: Mapping (dis) agreement in hydrologic projections, Hydrology and Earth System Sciences, 22, 1775–1791, https://doi.org/10.5194/hess-22-1775-2018, 2018.

Merleau, J., Perreault, L., Angers, J.-F., and Favre, A.-C.: Bayesian modeling of hydrographs, Water Resources Research, 43, W10 432, https://doi.org/10.1029/2006WR005376, 2007.

Middelkoop, H., Daamen, K., Gellens, D., Grabs, W., Kwadijk, J. C., Lang, H., Parmet, B. W., Schädler, B., Schulla, J., and Wilke, K.: Impact of climate change on hydrological regimes and water resources management in the Rhine basin, Climatic Change, 49, 105–128, https://doi.org/10.1023/A:1010784727448, 2001.

Milano, M., Reynard, E., Köplin, N., and Weingartner, R.: Climatic and anthropogenic changes in Western Switzerland: Impacts on water stress, Science of the Total Environment, 536, 12–24, https://doi.org/10.1016/j.scitotenv.2015.07.049, 2015.

Milner, A. M., Khamis, K., Battin, T. J., Brittain, J. E., Barrand, N. E., Füreder, L., Cauvy-Fraunié, S., Gíslason, G. M., Jacobsen, D., Hannah, D. M., Hodson, A. J., Hood, E., Lencioni, V., Ólafsson, J. S., Robinson, C. T., Tranter, M., and Brown, L. E.: Glacier shrinkage driving global changes in downstream systems, Proceedings of the National Academy of Sciences of the United States of America, 114, 9770–9778, https://doi.org/10.1073/pnas.1619807114, 2017.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. A., Mitchell, J. F. B., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P., and Wilbanks, T. J.: The next generation of scenarios for climate change research and assessment, Nature, 463, 747–756, https://doi.org/10.1038/nature08823, 2010.

Mote, P. W., Hamlet, A. F., Clark, M. P., and Lettenmaier, D. P.: Declining mountain snowpack in western North America, BAMS, pp. 39–49, https://doi.org/10.1175/BAMS-86-1-39, 2005.

Mote, P. W., Li, S., Lettenmaier, D. P., Xiao, M., and Engel, R.: Dramatic declines in snowpack in the western US, npj Climate and Atmospheric Science, 2, 1–6, https://doi.org/10.1038/s41612-018-0012-1, 2018.

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blodgett, D., Brekke, L., Arnold, J. R., Hopson, T., and Duan, Q.: Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: Data set characteristics and assessment of regional variability in hydrologic model performance, Hydrology and Earth System Sciences, 19, 209–223, https://doi.org/10.5194/hess-19-209-2015, 2015.

Oudin, L., Kay, A., Andréassian, V., and Perrin, C.: Are seemingly physically similar catchments truly hydrologically similar?, Water Resources Research, 46, W11 558, https://doi.org/10.1029/2009WR008887, 2010.

R Core Team: R: A language and environment for statistical computing, https://www.r-project.org/, 2019.

Ramsay, J., Wickham, H., Graves, S., and Hooker, G.: Package 'fda': Functional data analysis, https://cran.r-project.org/web/packages/fda/fda.pdf, 2014.

Ramsay, J. O. and Silverman, B. W.: Applied functional data analysis: methods and case studies, Springer, New York, https://doi.org/10.1007/b98886, 2002.

Rousseeuw, P. J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, Jounal of Computational and Applied Mathematics, 20, 53–65, 1987.

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, 2895–2911, https://doi.org/10.5194/hess-15-2895-2011, 2011.

Schlef, K. E., Moradkhani, H., and Lall, U.: Atmospheric circulation patterns associated with extreme United States floods identified via machine learning, Scientific Reports, 9, 1–12, https://doi.org/10.1038/s41598-019-43496-w, 2019.

Sharghi, E., Nourani, V., Soleimani, S., and Sadikoglu, F.: Application of different clustering approaches to hydroclimatological catchment regionalization in mountainous regions, a case study in Utah State, Journal of Mountain Science, 15, 461–484, https://doi.org/10.1007/s11629-017-4454-4, 2018.

Stewart, I. T., Cayan, D. R., and Dettinger, M. D.: Changes toward earlier streamflow timing across western North America, Journal of Climate, 18, 1136–1155, https://doi.org/10.1175/JCLI3321.1, 2005.

Tallaksen, L. M. and Hisdal, H.: Regional analysis of extreme streamflow drought duration and deficit volume, Friend'97 - Regional Hydrology: Concepts and Models for Sustainable Water Resource Management, 246, 141–150, https://doi.org/10.1212/WNL.0b013e31823ed0a4, 1997.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American Meteorological Society, 93, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

Ternynck, C., Ali, M., Alaya, B., Chebana, F., Dabo-Niang, S., and Ouarda, T. B. M. J.: Streamflow hydrograph classification using functional data analysis, American Meteorological Society, 17, 327–344, https://doi.org/10.1175/JHM-D-14-0200.1, 2016.

Thornton, P., Thornton, M., Mayer, B., Wilhelmi, N., Wei, Y., and Cook, R.: Daymet: daily surface weather on a 1 km grid for North America, 1980-2012, 2012.

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

USGS: USGS Water Data for the Nation, https://waterdata.usgs.gov/nwis, 2019.

Viviroli, D., Gurtz, J., and Zappa, M.: The hydrological modelling system PREVAH. Part II - Physical model description, Geographica Bernensia, 40, 1–89, 2007.

Vose, R., Easterling, D. R., Kunkel, K., and Wehner, M.: Temperature changes in the United States, in: Climate Science Special Report: Fourth National Climate Assessment, Volume I, edited by Wuebbles, D., Fahey, D., Hibbard, K., Dokken, D., Stewart, B., and Maycock, T., chap. 6, pp. 185–206, U.S. Global Change Research Program, Washington, DC, https://doi.org/10.7930/J0N29V45.U.S., 2017.

Wagener, T., Sivapalan, M., Troch, P., and Woods, R.: Catchment classification and hydrologic similarity, Geography Compass, 1, 901–931, 2007.

Ward, J. H.: Hierarchical grouping to optimize an objective function, Journal of the American Statistical Association, 58, 236–244, https://doi.org/10.1080/01621459.1963.10500845, 1963.

Wolock, D. M., Winter, T. C., and McMahon, G.: Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses, Environmental Management, 34, 71–88, https://doi.org/10.1007/s00267-003-5077-9, 2004.

Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D. P.: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, Climatic Change, 62, 189–216, https://doi.org/10.1023/B:CLIM.0000013685.99609.9e, 2004.

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, 4483–4498, https://doi.org/10.5194/hess-16-4483-2012, 2012.

Yevjevich, V.: An objective approach to definitions and investigations of continental hydrologic droughts, Tech. Rep. August, Colorado State University, Boulder CO, 1967.