

Reviewer comments – Comments to responses made by the author team

Paper entitled: "Understanding Changes in Iceland's Streamflow Dynamics in Response to Climate Change"

Journal: HESS

Authors: Helgason et al.

First Review submitted: 3 March 2025;

Comments to the author response submitted 19 September 2025

General comments – first round (*italics*)

including responses from authors (normal font) and reviewer's remark to these (blue).

The main aim is to “investigate the multi-annual variability in Iceland's climate and analyse changes in streamflow, aiming to link these changes to shifts in hydro-meteorological drivers and changes in glacier extent and mass”. The latter part “changes in glacier extent and mass” is not explicitly analysed in the paper through own analyses, rather the results are discussed in light of previously reported changes in glacier extent and mass. This in contradiction to shifts in hydro-meteorological drivers, which are quantified as trends, although the link to changes in streamflow is discussed in qualitative terms. This distinction should be made clear to the reader from the start.

While our study does not directly analyze glacier mass balance, we do compare trends in annual and summer melt-season streamflow to changes in glacier extent (Figure 10). Beyond this, glacier extent changes and mass loss are currently discussed in the context of previously documented trends rather than through independent calculations. We have made this clear to the reader from the start by making the following changes to the text: “[...] investigate the multi-annual variability in Iceland's climate and analyze changes in streamflow, aiming to link these changes to shifts in hydro-meteorological drivers, catchment attributes, and changes in glacier extent.”

Fine; however, it should be made clear in the text where the data on changes in glacier extent/area stem from, see Figure 9 (no reference included).

The study is thorough in that it investigates trends in streamflow and its meteorological drivers for different periods, temporal resolution and variables. The analysis is straightforward, using well-established methods. It is descriptive in its presentation, with an extensive result section, listing the results in a sequential way supported by a series of informative figures. As such, it resembles more a scientific report than a research paper. The results are discussed in a qualitative way, comparing the streamflow trends with trends in precipitation and temperature. The reasons for choices taken are not always well motivated; for example, what is the purpose/added value of looking at the daily trend in smoothed 21-day averages? And why is the spring freshet, rather than the timing of the peak flow (often used in the literature), chosen (or why not both)? Choosing indices commonly used in the literature, facilitates comparison with other studies. The paper would have benefitted from being tightened somewhat with a clear motivation for why - and how - the different analysis contributes to the aim and research question defined.

We acknowledge that the results section was extensive in the previous version of the manuscript. To enhance clarity and focus, we have streamlined this section by emphasizing the most critical findings that directly address our research questions. We have relocated detailed analyses (Figures 6 and 11 that show the sub-seasonal trends in meteorological forcings and streamflow 7 for all catchments) to the supplementary material. Since Figure 7 effectively summarizes these trends for meteorological forcings, we believe it is sufficient for the main text.

The paper has benefitted from moving material to the Supplement, although the presentation of the many trends is not always consistent when it comes to presenting non-significant trends vs significant trends. Sometimes a trend – encompassing only a few significant trends – are presented without a remark on significance. For example, related to Fig.5, it is written (author text in black):

- “The winter and spring seasons are characterized by modest positive or near zero trends”. These are non-significant trends and one should further not distinguish between modest and near zero trends as long as these are insignificant (nor is there a definition of ‘modest’ versus near zero’).
- “In the summer season, precipitation increased slightly in the northeast region and decreased in the southwest”. No comment on significance here either, although in this case there are a few significant trends.
- “Analysis of temperature trends for period 2 (Figure 5c) shows less statistical significance in seasonal temperature trends compared to period 1, with winter and summer exhibiting the least pronounced (and mostly insignificant) trends”. Summer does show a considerable number of significant trends in the southwest (although the number of significant vs non-significant trends is not provided).

These examples highlight the need for the authors to carefully go through the text and ensure that trends are presented in a consistent and scientific way. Although focus should be on whether a trend is significant or not, one may comment on regional consistency in trend direction even though these are not all significant, but this should then be made clear.

Further, any comment on magnitude on trends that are not significant should be made with care and preferable avoided, as these are just that – non-significant changes.

The combination of maps showing spatial trends in streamflow for different periods and seasons (shown as point values at the gauge, Figure 7) and heatmaps showing an overview of positive and negative trends for annual as well as seasonal trends in streamflow (Figure 8), is informative and a good starting point for discussion. Note, such heatmaps are not provided for meteorological forcings although stated in the replay above.

We have added a paragraph to the methods chapter that explicitly articulates the motivation for including sub-seasonal analyses. “While annual and seasonal trend analyses are valuable for identifying long-term hydrological changes, they may overlook important shifts occurring at finer temporal scales. Sub-seasonal trends can provide valuable insights even when no significant annual or seasonal trends are detected. Moreover, when seasonal trends are present, this analysis helps pinpoint when changes occur in greater detail.”

Fine, however, when commenting on September 20, the period this date represents given a 21-day moving average, i.e., 10-30 September, should be added to the text (e.g. in brackets).

For the spring freshet timing, our choice was based on its hydrological significance in cold regions, where the onset of snowmelt-driven runoff is a key indicator of climate-driven changes in seasonal hydrology. However, we recognize that peak flow timing is also commonly used in the literature, and we have added this analysis to the manuscript.

Fine.

The abstract states “We then analyze trends during the last 30 and 50 years (actually 31 and 51 years given that the start and end year are included) in annual, seasonal, and smoothed daily streamflow volumes, the timing of the spring freshet, and extreme flow conditions, linking these changes to environmental conditions and catchment attributes.” This latter is not entirely true as only three catchment attributes are listed (Table S3), one being the degree of anthropogenic impact (classified into four groups), which is used merely to exclude (heavy influenced) catchments from the study, not in the analysis itself. Of the other two, only the percentage of glaciers is included in the analysis and discussion, whereas BFI is not used at all. Accordingly, a quantitative analysis on how these attributes relate to the trends detected is missing. The streamflow series stem from a dataset (LamaH-Ice), which includes a wide range of catchments attributes, thus this is a missed opportunity. Including such an analysis would have enhanced the relevance and quality of the paper, allowing an in-depth analysis of the patterns found.

Our analysis is based on hydrological years, spanning October 1, 1973, and October 1, 1993, to September 30, 2023, which correspond to 50 and 30 complete years, respectively. We have clarified this in the manuscript (Sect. 2.6) to avoid ambiguity.

Fine.

We appreciate the reviewer’s suggestion regarding the inclusion of additional catchment attributes in the analysis. We have included an analysis of correlations of trends with all catchment variables in the LamaH-Ice dataset.

Fine; however, these should be introduced under Data, i.e., the specific catchment attributes included (e.g., as a table), and can be grouped into types in the main text for space reasons.

The exception being a separate analysis found in the Supplement. Figure S21 nicely shows the relationship between the calculated trend in spring freshet timing, the catchment mean elevation and the catchment average trend in spring (MAM) temperature. It is here concluded (period 2) that there is “a pattern of an increasing negative trend in spring freshet timing with higher catchment mean elevation”. Is this so? Figure S21 shows that the largest negative trends in the spring freshet timing (approx. -18 days/decade) is found at the lowest elevation (< 200m asl). These results are in either case not discussed or commented on later in the paper.

We appreciate the reviewer’s careful assessment of Figure S21 and their attention to detail. The figure illustrates the relationship between spring freshet timing, catchment mean elevation, and spring (MAM) temperature trends. In the manuscript, we stated: “At the highest elevations, we note a pattern of an increasing negative trend in spring freshet timing with higher catchment mean elevation.” This observation was based on the seven catchments with mean elevations above 800 m a.s.l. We acknowledge that the small number of high-

elevation catchments does not provide sufficient evidence to support a general conclusion of an increasing negative trend in spring freshet timing with higher elevation.

A more accurate observation is that there is no clear evidence that catchments at lower elevations experience systematically earlier freshet timing compared to higher-elevation catchments.

Agree, particular so if we only look at (the four) significant trends.

We have expanded this analysis to other catchment attributes. We have thus removed figure S21 and added a discussion on significant correlations between timing metric trends and catchment attributes. Interestingly, while no correlation is found between the timing of onset of spring freshet and catchment mean elevation, the timing trends are correlated with catchment elevation standard deviation, indicating that flatter catchments experienced an acceleration in snowmelt onset.

Good to include all catchment attributes, and these are interesting results. However, hard to judge their relative importance without knowing their co-correlation.

The discussion section is nicely written summarising key findings of the study, albeit rather brief when it comes to discussing the results as such. It would have benefitted from an overall broader discussion, including a more extensive section on how the findings compare to previous studies (from Iceland or comparable sites). Section 4.2.3 mentions the lack of other studies on Iceland, and refers to two publications (from 2006 and 2008). However, the European study by Blöschl et al., 2017 (mentioned in the Introduction) and the Nordic study by Wilson et al., 2010 (not in the reference list) both include Iceland (there may be others). How do the findings from these studies compare to the results of this study? It may be valuable to compare also with studies from similar regions in other parts of the world where glaciers are present (Section 4.2.3 does not mention glacier fed rivers), i.e., do the findings from this study agree with other studies or are they unique to Iceland?

We have made changes to the discussion section to compare our findings to previous studies. In line 54, after the sentence: “A key takeaway was that despite a substantial increase in precipitation, no corresponding increase was found in the flow of non-glacial rivers (Jónsdóttir et al., 2006).” we have added: “Wilson et al. (2010) analyzed trends in annual and seasonal streamflow in Icelandic rivers for 1961–2000 and found no significant trends in either annual or seasonal streamflow, confirming the findings by Jónsdóttir et al. (2006). Their analysis also showed no shift in the timing of the spring floods.”

In the discussion section (line 507), we have added: “Our analysis shows that the timing of the spring freshet has not systematically shifted, consistent with the findings of Wilson et al. (2010) for the period 1961–2000.”

Fine.

Blöschl et al. (2017) found that the timing of annual maximum floods in southwestern Iceland shifted later in the year during the period 1960–2010, while in northeastern Iceland, flood timing remained stable or occurred earlier, as noted in the introduction. Our results show that the timing of annual peak flows has not systematically shifted in the over the periods that we analyzed.

This reference is now included in the Introduction, but not referred to in the Discussion. The latter is recommended as the key point of mentioning related studies are to compare their findings with your study.

We have also expanded the discussion section to discuss findings in the analysis of trends in streamflow variability (coefficient of variation, flashiness index, baseflow index) and correlations between streamflow trends and catchment attributes.

Fine; suggest to link these analyses to the first part of the study, i.e., multiannual variability in hydrometeorological variables (rather than after the trend analysis). It can then be picked up upon when discussion the trends.

To maintain focus and conciseness, we will not expand the discussion to include comparisons with studies from other glaciated regions.

This does not appear as a valid argument, notable as it is argued that the study contributes to global knowledge on ... (last sentence in abstract). Thus, it is recommended to link the results to similar studies in comparable environments. This does not have to be comprehensive (or a global review), but it is important to put the results in a wider context, particular as it provides mixed results as to trend magnitude and direction (variability in both space and time).

The paper has an extensive Supplement assessing the Homogeneity of streamflow records (S1), Changes in evapotranspiration compared to changes in precipitation (S2), Trends in streamflow (S3) and Overview of gauges used in the study (S4). Most of this material are not considered necessary to include (see suggestions below).

Fine, the Supplement has undergone major changes.

In Conclusion: Major revisions are recommended. Accordingly, detailed textual comments are not provided at this stage, but general recommendations provided.

Issues that need clarification/discussion

1. Presentation of the dataset:

- *It is stated that 38 stations (later reduced to 37 due to the homogeneity test) were used, however, in Section 3.2 one refers to the whole dataset of 107 catchments in the LamaH-Ice dataset, including human influenced catchments, and in the Data section reference is made to 79 stations (Section 2.3). This is confusing;*

We analyze trends in meteorological forcings using data from all 107 catchments in the LamaH-Ice dataset. For streamflow trends, we use 25 gauges for period 1 and 37 gauges for period 2 (following the homogeneity test). We acknowledge that the current wording may be unclear, and we have revised Section 2.3 to explicitly state the number of stations used in each case to avoid confusion

It is now clearer which number of stations that are used, but Section 2.3 still states “In this study, we use timeseries for total precipitation, snowfall, and temperature

for all 107 catchments in LamaH-Ice.” This still appears confusing, and if the case, it should be clearly stated for which analysis these 107 stations are used. Further, related to Figure S2, one refers to 111 catchments.

- *Are any of the catchments analysed sub-catchments of larger catchments included?*

A few of the catchments included in our analysis are sub-catchments of larger catchments that are also analyzed, though this is relatively uncommon. This can be observed in the figures displaying basin outlines along with gauge locations (e.g., Figures 3, 8, 12, and 13). To ensure clarity, we have explicitly stated this in the Data section, as follows: “Some gauges represent nested sub-catchments of larger river systems. Specifically, gauges 46 (Jökulsá á Fjöllum at Upptyppingar) and 59 (Kreppa) are sub-catchments of gauge 45 (Jökulsá á Fjöllum at Grímsstaðir); gauge 86 (Tungnaá at Maríufoss) is nested within gauge 102 (Þjórsá at Þjórsártún); and gauges 8 (Brúará at Dynjandi), 36 (Hvítá at Fremstaver), and 79 (Sog at Ásgarður) are sub-catchments of gauge 98 (Ölfusá at Selfoss). Rather than exclude gauges 45, 98, and 102, we retain them in the analysis because they integrate substantial additional drainage areas downstream of their respective sub-catchments.”

This is an important clarification, and needs to be reflected in the discussion, i.e., these three stations do not represent independent stations when it comes to number of stations with significant trends. On the other side, the trends of sub-catchments can provide important information about spatial variability in trends, e.g., whether they are consistent or diverse.

- *The analysis is done for two different periods; 1973-2023 and 1993-2023, 31 and 51 years respectively). It would have been interesting to know the mean streamflow, precipitation and temperature for the two periods.*

We have added a table in the Supplement (Table S2) that shows the mean streamflow, precipitation, and temperature for the two analyzed periods.

Fine; however, this is not reflected in the discussion when summarised over the stations, i.e., is Period 1 overall wetter/drier and warmer/cooler than Period 2? Are there regional differences among stations?

- *In Figure 1 the period 1951-2024 is used; why is this period chosen? Later in the text, the year 2024 is highlighted as a particular dry year and has been given a separate paragraph in the discussion. Why this focus on 2024 if not included in the trend analysis?*

The period 1951–2024 was chosen primarily based on data availability from the ERA5- Land reanalysis. Initially, ERA5-Land was available from January 1, 1950, but it has since been extended back to 1940. In the updated manuscript we now use the period 1950-10-01 to 2024-09-30. The year 2024 is highlighted

separately in the discussion due to its extreme conditions, which underscore the high variability in precipitation and temperature in Iceland. Although 2024 is not included in the trend analysis due to the lack of available streamflow data for that year, its recent occurrence makes it particularly relevant for water management and stakeholders in Iceland. We have clarified this distinction in the revised manuscript.

Fine; However, it should be introduced up-front in the paper that the 2024 event will be discussed in particular. Now this appear a bit ad-hoc in the discussion.

- *Figure 1 further includes a rolling 5-year mean and trendlines; how do these trendlines relate to the trends later calculated?*

The trendlines in Figure 1 provide a visual representation of overall long-term changes, while the formal trend analysis later in the manuscript is based on different time periods. We have explained this difference

Fine; however, it is important to state whether the trendlines shown are derived based on annual values or the rolling 5-year mean. The latter is not recommended due to dependencies among the data points.

- *In the first part of the study (annual variability) the hydrological year is used; is this also the case for the second part of the study on trends?*

The hydrological year is used for both annual variability and trend analysis. We have confirmed this explicitly in the text: In sect. 2.6, after the sentence: “We calculate trends for annual, seasonal and sub-seasonal averages.” We have added: “Annual trends are based on hydrological years, defined as the period from October 1 to September 30.”

Fine.

- *The seasonal study includes four commonly used seasons where summer is defined as June, July, August. Later, another summer season is introduced, the i.e., July, August, September also referred to as the summer melt season, as being more relevant for assessing the peak contribution from glaciers. This choice ought to be well introduced earlier (now it appears for the first time in the Result section) and discussed in terms of interpreting the results of the two different ‘summer’ periods.*

The introduction of two different summer definitions (JJA and JAS) has been clarified in chapter 2.6, along with the rationale for considering JAS as a more relevant period for glacier melt contributions. We have added: “In addition to the summer season, we define a glacial melt season as July to September to better isolate runoff originating from glacier ablation. While the conventional summer period includes both snowmelt and glacier melt, the July to September window

reduces the influence of seasonal snowmelt. This separation allows for clearer interpretation of glacier specific contributions.”

Fine.

- *The area of the catchments should be provided as it is important for evaluating the spatial variability in meteorological variables, notable in large catchments where a mixed trend pattern may impact the catchment average trend signal;*

We have included catchment areas in Table S1 in the Supplement to help assess spatial variability in meteorological trends.

Fine; however, I miss a comment on the size of the catchments when discussing trends, as a large catchment may experience diverging trends across the catchment area. This can be assessed by looking at trends in gridded ERA5-Land (see link below to the Climate Change Knowledge Portal).

- *Uncertainty is not discussed, neither in terms of the observed data (streamflow) nor in the gridded ERA5-Land variables (temperature and precipitation), and how this may impact the results;*

We recognize the need to discuss data uncertainties. We have added the following sections to the manuscript: “The uncertainty in the streamflow observations is discussed in detail in the LamaH-Ice data description paper (Helgason and Nijssen, 2024a). Streamflow measurements in Iceland are prone to interruptions (e.g., ice disturbances or instrument malfunctions), particularly during winter, which can reduce data availability and introduce additional uncertainty. Moreover, uncertainty in older streamflow periods is higher than in recent periods due to older instrumentation with greater uncertainty.”

Fine; however, based on this assessment, would you consider some of the results more uncertain than others? Ref. winter streamflow measurements being more prone to interruptions than summer flow and the fact that the main low flow season is in winter.

- *It is stated that precipitation in the ERA5-Land reanalysis is underestimated for Iceland (Helgason and Nijssen, 2024a); what about temperature? Cryosphere processes are very sensitive to air temperature, notable the zero-degree crossing.*

To address this, we have assessed temperature differences between ERA5-Land and another regional reanalysis (RAV-II) for the catchments in LamaH-Ice. See discussion in Sect. 2.3 in the main manuscript and figure S2 in the Supplement.

This provides additional information, but the comparison of two reanalysis datasets does not necessarily contain information about the representativity of ERA5-Land versus observations. The revised Supplement compares observed temperature for a weather station (Reykjavik), with catchment average temperature derived from ERA5-Land. Why not compare the grid average of the

grid cell where the weather station is located? This would provide a more direct comparison.

It is stated that ERA5-Land is slightly cooler than the observed series. Is a difference around 2.5 °C seen as 'slight'?

Further, in the presentation of Figure S3, it is said that "both reanalyses indicate increasing precipitation trends in the east and northeast, transitioning to decreasing or near-zero trends toward the southwest. This spatial coherence strengthens confidence in the qualitative pattern of change." Looking at the maps, there are clear differences in the trends, also in the direction of trends (some of which are significant, in both increasing and decreasing direction). This should be corrected in the description and implications for the study discussed.

- *How does ERA5-Land perform for other glaciated areas of the world?*

While ERA5-Land's global performance is an interesting question, our study is focused specifically on Icelandic conditions.

The comment related in particular to glaciated areas and whether there were any studies evaluating the performance for these areas. Recommended to check.

- *Has the ERA5-Land dataset been evaluated for its ability to reproduce trends in observations for Iceland? If not, it is recommended to include an at-site comparison where observed (station) precipitation (P) and temperature (T) time series are compared with the ERA5-Land grid cell representing the location of the stations;*

ERA5-Land has not been explicitly evaluated for trend reproduction in Iceland. This has been done for temperature in other areas, e.g. Turkey (Yilmaz, 2023) and China (Zhao & He, 2022). We have included a comparison of observed station-based temperature time series with ERA5-Land data (See discussion in Sect. 2.3 in the main manuscript and figure S1 in the Supplement). Precipitation observations are difficult to use for this kind of comparison due to snow and wind in winters. However, we have compared trends in precipitation between ERA5-Land and another reanalysis dataset (CARRA: See discussion in Sect. 2.3 in the main manuscript and figure S3 in the Supplement).

See comments above; the obvious would be to compare trends based on station data with trends in the corresponding grid cells in ERA5-Land.

- *ET values are likely rather low on Iceland, a small change in ET (in mm), may become a notable percentage change (add to the discussion);*

Evapotranspiration (ET) values are indeed low in Iceland. Due to this, even small absolute changes in ET (measured in mm) can correspond to significant relative

percentage changes. This is why we report ET trends as a percentage of the average watershed precipitation per decade in Figure 4.

Fine.

- *It is suggested to present spatial trend patterns in the meteorological variables for the whole of Iceland. By mapping the trends for each grid cell (ERA5-Land dataset) over Iceland, one can detect potentially regional diverging trend pattern that may help understand the spatial aggregated trends in streamflow for large catchments (as they may cover an area with mixed trend patterns).*

We analyze trends for all 107 catchments in LamaH-Ice, providing high spatial coverage across Iceland. Given the consistency of trends observed across regions (see Figures 4 and 5), a gridded analysis would not add substantial new insights. Our approach ensures that catchment-scale trends, which are most relevant for streamflow analysis, are well represented.

The Climate Change Knowledge Portal explores trends in temperature and precipitation for different periods and seasons on a gridded as well as catchment scale for countries around the world, including Iceland. Thus, these data exist, allowing to detect regional as well as within catchment spatial variability in trends. It can be highlighted that the current paper adds to existing knowledge by exploring catchment average trends and the link to streamflow trends.

[Iceland - Trends & Variability \(ERA5\) | Climate Change Knowledge Portal](#)

2. Supplementary material

[Comments have been satisfactory accounted for](#)

3. Aggregating climate trends across the catchment area

[Comments have been satisfactory accounted for](#)

4. Calculation and discussion of trends

[Remaining comments \(those resolved have been removed\)](#)

- *Further, if more than 10% are missing, that year is excluded from the trend analysis. How is a missing year dealt with in the time series; just skipped or indicated as missing? Is the assessment of significance adjusted accordingly?*

When a given year (or season) has more than 10% missing daily data, we exclude that year from the analysis. This ensures that only years with sufficiently complete records are used to represent the typical flow conditions. We omit series with more than 20% of annual/seasonal values missing. A missing year is dropped from the series before calculating trend and significance. We acknowledge that omitting these years reduces the total number of observations and may affect the

trend estimates and the significance test's power. This is discussed in Sect. 2.3 (Data)

For clarification: did you mark the year as missing in the time series prior to the trend analysis or simply remove it (resulting in a shorter time series)?

- *The trend direction is shown as blue and red. Is there a range in z-values for when a trend is defined as zero?*

We do not use a range in z-values for zero trends.

Zero trend must still be indicated through the accuracy in the z-value. The question further relates to how to report positive and negative trends, and restrain for commenting on weak and non-significant trends.

- *Suggest to focus on discussing significant trends and only comment on non-significant trends if there are a clear regional or temporal pattern, which then contains information beyond the individual station.*

Our manuscript already follows this approach for meteorological trends, where we primarily discuss significant trends and comment on non-significant trends only when they exhibit a clear regional or temporal pattern. To ensure consistency across all variables, we have refined the results and discussion sections to apply this approach to streamflow trends as well.

Fine; however, there are still inconsistencies as to how non-significant trends are presented, ref. earlier comments.

- *Low flow and high flow are represented by the 10th and 90th percentile from the whole time series, and no distinction is made between seasons. As the dominating low flow or high flow season may be in summer or winter (or both), events generated by different processes may be included in a given time series, which hampers a discussion on the causes of the trend seen.*

We have distinguished seasonal contributions to high and low flows to improve the interpretation of percentile-based streamflow trends.

Fine.

5. Streamflow anomalies

Comments have been satisfactory accounted for

6. Sub-seasonal (daily) trends

One may argue that a choice of a 30 (or 31) day moving window may be better than a 21-day window, as it resembles the monthly time resolution (this is also the time window used for gap filling).

We tested multiple window lengths and found that a 21-day window effectively balances the need for temporal precision while maintaining robustness in the analysis.

Fine.

A miss a discussion on the cause of the sharp, and consistent, transition between positive and negative trends seen in Fig. 6.

Miss an answer here.

The overlying dashed black line is hard to see in the figures.

We have enhanced the visibility of the dashed black line indicating trend significance.

Fine.

7. Glacier catchments

Trends in streamflow are presented for all catchments in the main body of text, whereas the results for glacier catchments are presented in the Supplement (S3). Why not present the results for non-glaciated catchments as well?

Figure S20 in the Supplement presents trends in annual and summer melt-season streamflow, specifically for glacier-fed rivers. This figure facilitates a direct comparison between annual and melt-season trends in glacial rivers. In contrast, the main manuscript presents the trends in annual and summer streamflow for all gauges, ensuring a broader overview.

Fine; (should be Figure S7).

Consistency in terminology:

- *Both watershed (US) and catchment (UK) are used (choose one)*
- *Both fall (US) and autumn (UK) are used (choose one)*

Both terms are understood and accepted in both countries and widely used in English around the world. The interchangeable use of these terms does not lead to confusion and we therefore see no need to change this.

I disagree and suggest to be consistent. It was further noticed that also the term 'basin' is used; i.e., three terms are used for the same term (number of times given in bracket): basin (8), watershed (7) and catchment (49). What to gain by using different terms, is not clear.

Reviewer comments – second round

Paper entitled: "Understanding Changes in Iceland's Streamflow Dynamics in Response to Climate Change"

Journal: HESS

Authors: Helgason et al.

Second review submitted 19 September 2025

General comments

Overall, the paper has significantly improved, responding in a thorough way to the many questions and remarks made. It resembles more a paper now, still there are room for improvements, particularly related to a clearer structure (in what material is presented where), and the need to consistently and carefully present and interpret significant versus non-significant results. Related to the first of these, the data and the methods should be introduced up front, avoiding introducing new data or methods in the Result section. The authors are also recommended to better link among the different results in the Discussion section (some examples given below).

Overall, the Discussion section is nicely written and emphasis is given to key results. The authors should ensure that any outcomes and updates following this second revision is evidenced in the Discussion and Conclusions. Further, it is recommended to reflect on the many uncertainties that may affect the results, as well as provide a wider perspective by referring to similar work in other similar environments outside Iceland.

The comments below supplement the response given to the first round of comments (see separate file).

1. Overall, choices made are not always well argued, it is often stated this is what we do, not why.
2. Carefully check Figure legends as they:
 - do not always refer to the subfigures (e.g. a), b) and c)).
 - do not need to repeat methodology in the legend (ref. Figure 1. – how trends are calculated and Fig 3 – how correlations are calculated).
 - sometimes repeats heading of subfigures and use different fonts, e.g., Figure 4 and Figure S9 (periods are given at three places).
3. The result section separates between multiannual variability (Section 3.1) and trend analysis (Section 3.2); however, trends are also presented in 3.1.1. Thus, the result section should be better structured.

Specific comments

- Abstract: sub-seasonal trends are not mentioned here.
- Section 2.2: the heading is misleading as it does not describe the hydrological regime of Iceland (which should be introduced); rather it focuses on the climate variability and its link to large scale atmospheric patterns.

- Data: Should add how many stations contains > 10% missing values. This can be added to the Table S1 (important as it says something about the uncertainty in the trend analysis for specific stations).
- Data: It is not sufficient to state a ‘selection of catchment attributes’ without stating what these are (if one does not want to list all, one can group these, e.g. soil parameters) and referring to where these are described in details.
- It is vital to know how the BFI was calculated. Given that it is a hydrograph separation method, it will be affected among other by the climate, e.g. catchment in wet climates (or during a wet period) will generally have a higher BFI than a similar catchment in a dry climate (or during a dry period), simply because of higher flows and the lack of longer dry periods.
- Section 2.6: the trend method is not insensitive to outliers, rather it is less sensitive than other methods (e.g. linear regression).
- Section 3.1.1 under Results. The first two sentences report from literature and should thus go to the Introduction.
- Section 3.1.1 Changes in CV, flashiness and BFI are recommended to move to this heading (multi-annual variability).
- Figure 1: Snowfall anomalies are shown; is the change in the contribution of snowfall vs rainfall quantified? Ref. statement line 259-261.
- Figure 1: The text refers to low flow periods in the annual plots; this should rather be lower than normal annual flows (low flow periods mean something else).
- Figure 1: Enhanced glacier melt around 2010 is commented on; how is this explaining the patterns seen.
- Figure 2: Valuable that the authors have included the percentage of glacier in the anomaly plot. The figure shows that glaciated catchments show a consistent reduction in annual streamflow (glacier % >10), with higher anomalies for the highest glaciated catchments. How does this agree with the trend results seen in S5 (albeit different periods)?
- Looking at causal factors (link large scale atmospheric patterns) for the annual variability is not introduced as an aim for the study, and part of the text in Section 3.1.3 relates to Data and Method and should be moved accordingly.
- Section 3.1.3: What is the common factor for catchments showing high correlations?
- Section 3.2: here it is referred to 107 catchments, but these are not included in the results (see earlier comment). Further, information about the data used are repeated.
- Figure 6: Are these trends calculated based on the grid averages for the whole of Iceland?
- Figure 6: Make sure to focus the presentation/discussion on significant trends, emphasising that the warming signal is more pronounced in the longer period.
- Figure 8: this figure nicely summarises the trends in streamflow; it would be valuable to see a similar figure for temperature and precipitation as it is often stated that there is a higher number of positive/negative trends without quantifying this.
- Line 389; why would positive trends in spring streamflow (i.e., increasing flow) correlate negatively with temperature and total ET, implying that higher Q is linked to lower temperature and lower ET? It is suggested by the authors that this is linked to higher evaporative demand causing reduced flow, however Q is increasing in this case. Please clarify.

- Higher temperature may both lead to an increase (snow/glacier melt) and a decrease in streamflow (higher ET). The discussion should better balance this.
- Figure 9: it is not commented on which of these trends are significant.
- Changes in some flow indices show a positive relation to BFI (e.g. summer streamflow), some a negative (e.g. spring streamflow); is this well explained?
- Figure 10: nice to include also the trend in peak streamflow, albeit the results are very diverse across regions.
- You may consider reducing the use of 'in this study' in the text.
- The correlation heatmaps in the Supplement lack reference to significance.
- The correlation between the variables should preferably be shown in a matrix, and perhaps the many heatmaps could be summarised in a correlation matrix (or two) as well.