

Manuscript egusphere-2024-4186: Response to reviewer

Hörður B. Helgason, Andri Gunnarsson, Óli G. B. Sveinsson and Bart Nijssen, Understanding Changes in Iceland's Streamflow Dynamics in Response to Climate Change

We would like to sincerely thank the reviewer for their exceptionally detailed and constructive feedback, which has greatly improved the scientific rigor, clarity, and structure of the manuscript. We have carefully revised the paper in response to all comments and implemented substantial edits throughout, including a reorganization of sections for improved readability and a consistent application of the significance-first reporting convention when presenting statistical results. The manuscript has also been refined to ensure that all data, methods, and interpretations are introduced and discussed in a coherent and logically structured manner.

In the following pages, the authors provide detailed, point-by-point responses to each comment from both rounds of review.

First round of major revisions:

- Reviewer comments (*italics*)
- Author responses (normal font)

Second round of major revisions:

- Reviewer comments (blue)
- Author responses (red).

Remaining comments (those resolved in the first round of revisions have been removed).

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)

We have added text in Sect. 2.3 Data to explain where the changes in glacier extent stem from.

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 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.

We thank the reviewer for the valuable comments regarding the consistency of language when reporting significant and non-significant trends. In response, we have thoroughly revised the manuscript to apply a clear *significance-first* reporting convention throughout the Abstract, Results, Discussion, and Conclusions. We believe these revisions substantially improve the clarity and scientific rigor of the manuscript and fully address the reviewer’s concern about consistent communication of statistical significance.

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 heatmaps for trends in meteorological forcings to the Supplement (Figure S1).

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)

We have added this to the figure caption and to the text.

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.

We have added this text to Data:

“We incorporate static catchment attributes from the LamaH-Ice dataset, which provide information on topography, climate, and hydrology, as well as land cover, vegetation, soils, geology, and glaciation. In addition, we use time series that describe temporal changes in glacier-covered areas within the catchments. Further details are available in the LamaH-Ice publication (Helgason and Nijssen, 2024).”

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.

We have added correlation matrices for the catchment attributes in the Supplement (Figures S14-18).

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 added this to the discussion section.

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.

Thank you for this suggestion. We considered integrating this material into Section 4.1 on multiannual variability. However, that section addresses variability driven by large-scale atmospheric and glaciological influences (e.g., the AO, SST anomalies, glacier mass balance) on multi-decadal time scales. In contrast, the coefficient of variation, flashiness index, baseflow index, and their relationships with catchment characteristics are directly tied to the trend analysis and describe changes in intra-annual flow behavior rather than decadal climate variability.

To address your suggestion, we included a linking statement in the transition between Sections 4.1 and 4.2, making explicit that the observed multiannual variability provides important context for understanding the trends in both mean flow and variability metrics. We believe this preserves a logical structure while responding to your request to link the two parts of the discussion.

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).

We agree with this comment, and we have added a paragraph at the end of Sect. 4.2.3 to put our results into context with results from other glacierized regions.

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.

To improve clarity, we have revised the sentence in Section 2.3 to: “In this study, we use timeseries for total precipitation, snowfall, and temperature for the catchments in LamaH-Ice.”. To ensure the distinction is explicit, we have also added the following sentence at the start of Section 3.2: “Note that trends in streamflow (Sect. 3.3) are only calculated for a selection of catchments with sufficient streamflow data coverage (Sect. 2.3).”

Regarding the reference to 111 catchments, we have updated the Supplement to consistently use 107. LamaH-Ice includes 107 unique gauges. The higher count of 111 results from four duplicate catchments created due to permanent diversions that altered the effective contributing areas of certain gauges. For clarity and consistency, we base our analyses on the 107 unique 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 Upptýppingar) 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 subcatchments.“

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.

Thank you for this observation. We have revised the sentence in Section 2.3 to reflect your point more clearly:

“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, and trends in the nested catchments can provide insight into spatial variability.”

To address the issue of non-independence, we examined the nested catchments separately. Only one of the sub-catchments show statistically significant streamflow trends (Jökulsá á Fjöllum, 1973-2023, SON and JAS seasons). We have clarified this in the figure captions (Figures 7 and S7) and in the text. Other than that, inclusion of subcatchments does not affect the counts of gauges with significant trends or the overall interpretation of the results.

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?

We have added an analysis of this in the supplement (Figure S23) and referenced this in the main manuscript (Sect. 2.3).

- 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.

Thank you for pointing this out. We agree that the 2024 event should not appear as an isolated remark in the Discussion. In fact, to shorten the already lengthy paper, we have removed the discussion of the 2024 year. Since meteorological data for the 2024–2025 water year are now available, we have extended Figure 1 to include this period and updated both the Results and Discussion sections accordingly. To address your comment, we now state in the Results section:

“Although these two years are excluded from the trend analysis in later sections due to incomplete streamflow records, their contrasting hydroclimatic conditions provide a contemporary example of the variability discussed here.”

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.

The trendlines shown are based on annual values, we have now explicitly stated that in the figure caption.

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 discussion 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).

In Sect. 3.2 we have added the following sentence: Note that meteorological forcings are calculated as catchment averages and may smooth over locally diverging trends within large catchments.

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.

To address this comment, we have added the following sentence to Sect. 2.3: “As a result, trend estimates that rely heavily on winter flows or on early parts of the record (e.g. the 1970s and 1980s) should be interpreted with greater caution than those based on recent summer flows.”

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 a 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 contains 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.

Thank you for the suggestion. We agree that grid-cell-to-station comparisons provide a direct assessment of ERA5-Land’s representativeness at specific locations. However, both the ERA5 grid cell and the catchment average represent area-averaged conditions over spatial scales much larger than a single station. As such, neither is expected to reproduce point-scale measurements exactly. Our focus was therefore on catchment-average consistency, which is most relevant for hydrological analysis.

It is stated that ERA5-Land is slightly cooler that the observed series. Is a different around 2.5 °C seen as ‘slight’?

We have removed the word “slightly” in the text. While the elevation difference between the Reykjavik station (60 m a.s.l.) and the catchment mean (171 m a.s.l.) partly explains the temperature offset, it cannot account for the full 2.5 °C difference. Assuming a typical moist environmental lapse rate of about 0.6 °C per 100 m, the expected elevation-related temperature difference would be approximately 0.7 °C. The remaining bias could reflect the distance between the weather station and the catchment.

This comparison shows that ERA5-Land can reproduce trends in measured temperature.

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.

We thank the reviewer for this comment. We have now updated this figure (S4) with the most recent version of CARRA precipitation data from the LamaH-Ice dataset (Version 1.5, Helgason & Nijssen, 2025). Our investigation revealed that the original CARRA data we used contained erroneous values for 2022-2023, with precipitation systematically underestimated during this period (approximately 5-10% lower than both the updated CARRA data and ERA5-Land reanalysis). This data quality issue led to an underestimation of precipitation trends in our original analysis.

The updated figure now shows much better agreement between precipitation trends calculated from ERA5-Land and CARRA. Both datasets now consistently show increasing precipitation trends in the east and northeast, transitioning to decreasing or near-zero trends toward the southwest. This improved consistency strengthens our confidence in the observed precipitation changes and their relationship to streamflow trend

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.

We have added a summary in Sect. 2.3.2 noting that ERA5-Land has been shown to represent temperature trends well in glacierized regions of Russia, the Tibetan Plateau and the Andes, while precipitation and snow-related variables are more uncertain and best used in a relative rather than absolute sense.

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

We appreciate the reviewer's comments and suggestions regarding the evaluation of ERA5-Land temperature and precipitation for Iceland. In addition to the temperature station-based comparison presented in the revised Supplement (Fig. S2), we have also assessed temperature differences between ERA5-Land and the regional reanalysis CARRA for all LamaH-Ice catchments (Sect. 2.3, Fig. S3), showing a close fit.

This dual comparison provides complementary perspectives: the CARRA assessment highlights large-scale spatial consistency in temperature fields across Iceland, while the station-based analysis offers a localized evaluation against observations to show that observed long-term trends are reproduced in ERA5-Land.

For precipitation, direct measurement comparisons are inherently uncertain, especially in winter, due to gauge undercatch in high winds and low temperatures. To address this, we compared precipitation trends between ERA5-Land and CARRA, which showed a close overall agreement. CARRA has been extensively evaluated against in situ measurements in Iceland, demonstrating good agreement for both temperature and precipitation (Rögnvaldsson et al., 2025). In that evaluation, precipitation data were filtered to include only observations under wind speeds below 8 m s^{-1} and temperatures above $2 \text{ }^{\circ}\text{C}$, thereby minimizing measurement biases. This supports our use of CARRA as a reliable reference dataset for assessing ERA5-Land performance. Additional details are provided in the Supplement (Sect. S2.2).

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](#)

- 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)?

We removed such years entirely from the time series rather than keeping them as missing values (e.g. NaNs or placeholders). Thus, the trend and corresponding p-value were computed only on the remaining valid years, without any gaps or empty timestamps influencing the calculation.

- 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.

We thank the reviewer for raising this important point. In our analysis, we did not apply a range in z-values to define a “zero” trend. Instead, all trends were derived directly from the Mann–Kendall test and Theil–Sen slope estimates. The magnitude and sign of the z-value therefore reflect the direction and strength of the trend, while statistical significance is determined strictly by the $p < 0.05$ threshold.

In response to the reviewer's concern about overinterpreting weak or non-significant results, we have systematically revised the manuscript to ensure that only statistically significant trends are

interpreted in the main discussion and figures. Non-significant trends are mentioned only when consistent directional tendencies are evident across multiple catchments or regions, and they are explicitly described as non-significant. This approach ensures transparency in reporting while avoiding overstatement of weak 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.

We thank the reviewer for being persistent with this. See the answer to the previous comment.

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

We have added the following sentence in the text in Sect. 3.2.2: “The alternating positive/negative precipitation trend swings visible in Figure 6b during non-significant periods arise from the high variability in day-specific trend estimates. These fluctuations should not be interpreted as real regime shifts but rather as statistical noise around near-zero long-term change.”

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

Agreed that consistency improves readability. We have therefore replaced all instances of “basin” and “watershed” with “catchment” and standardized the seasonal terminology by replacing all instances of “autumn” with “fall.”

Reviewer comments – second round, 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 nonsignificant 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.

We thank the reviewer for this observation. We agree that Section 3.1.1 contains visual long-term trendlines, which could be confused with the formal statistical trend analysis introduced in Section 3.2. To clarify the structure, we have changed the subsection title to:

“3.1.1 Multiannual variability and long-term tendencies in temperature and precipitation.”

We also emphasize in the text that the trendlines in Figure 1 are included only as “a visual representation of overall long-term changes, while the trend analysis of streamflow and meteorological variables in sections 3.2 and 3.3 is based on different time periods.”

Specific comments

- Abstract: sub-seasonal trends are not mentioned here.

We agree that sub-seasonal trends are an important component of the results. Since the abstract must balance brevity with coverage, we chose to highlight only the most prominent sub-seasonal signal. In the revised abstract, we now explicitly refer to the September precipitation increase as the dominant sub-seasonal change.

- 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.

We thank the reviewer for this observation regarding Section 2.2. In response, we have renamed Section 2.1 to ‘The hydrology of Iceland’, as it describes the hydrological regime of the country. Section 2.2 is now titled ‘Climate variability and link to large-scale atmospheric patterns’ to more accurately reflect its focus on climatic controls. We believe these revised subsection titles more clearly represent the content of each section.

- 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).

We have added this to Table S1.

- 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.

We have now explicitly stated in the text which catchment attributes are reported in table S1. Further, we have added this text to address your comment: “We incorporate static catchment attributes from the LamaH-Ice dataset, which provide information on topography, climate, and hydrology, as well as land cover, vegetation, soils, geology, and glaciation. In addition, we use time series that describe temporal changes in glacier-covered areas within the catchments. Further details are available in the LamaH-Ice publication (Helgason and Nijssen, 2024).”

- 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.

We thank the reviewer for highlighting the need to specify how the baseflow index was calculated. We have now expanded Sect. 2.6 to include the full configuration of the Lyne and Hollick digital filter used ($\alpha = 0.925$, three passes, 30-day reflection) and clarified that the filter was applied to the full time series before aggregating to annual values to avoid edge effects.

We also note explicitly that the BFI is interpreted as a relative indicator of baseflow contribution, acknowledging that it may be influenced by both hydrological characteristics and climatic wetness conditions.

- Section 2.6: the trend method is not insensitive to outliers, rather it is less sensitive than other methods (e.g. linear regression).

We have updated the text to reflect this.

- Section 3.1.1 under Results. The first two sentences report from literature and should thus go to the Introduction.

We have moved this to the introduction.

- Section 3.1.1 Changes in CV, flashiness and BFI are recommended to move to this heading (multi-annual variability).

We considered integrating this material into Section 4.1 on multiannual variability. However, that section addresses variability driven by large-scale atmospheric and glaciological influences (e.g., the AO, SST anomalies, glacier mass balance) on multi-decadal time scales. In contrast, the coefficient of variation, flashiness index, baseflow index, and their relationships with catchment characteristics are directly tied to the trend analysis and describe changes in intra-annual flow behavior rather than decadal climate variability. We thus decided not to follow this suggestion.

- Figure 1: Snowfall anomalies are shown; is the change in the contribution of snowfall vs rainfall quantified? Ref. statement line 259-261.

We have now quantified this in the text as follows:

“The partitioning of precipitation into rainfall and snowfall (Figure 1c) shows that snowfall contributed substantially to annual totals during 1970–1995, averaging 54%. This declined modestly but significantly to 47% during 2000–2025, while the rainfall fraction increased from 46% to 53%. Although precipitation phase exhibits pronounced decadal-scale fluctuations since 1950, the rainfall fraction shows a significant long-term upward trend of 0.58% per decade (Mann–Kendall $p = 0.047$), indicating a gradual shift toward a rainfall-dominated precipitation regime.”

- 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).

Presumably this refers to figure 2. We have fixed this in the text.

- Figure 1: Enhanced glacier melt around 2010 is commented on; how is this explaining the patterns seen.

We assume this comment refers to figure 2 as well. We have added the following text: “This is evident in the elevated flows in 2010–2011 in highly glaciated rivers such as Jökulsá á Dal,

Jökulsá á Fjöllum, and Djúpa (noting that the 5-year rolling mean extends into subsequent years).”

- 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)?

We have added the following paragraph into Sect. 4.2.3: “For glaciated rivers, our results showed positive long-term streamflow trends (1973–2023) but negative or insignificant recent (summer) trends (1993–2023). This contrast reflects a transition from a high-melt period in the early 2000s to reduced glacier runoff after 2010. As illustrated in Figure 2, exceptionally high flows around 2000 elevate the long-term trend, whereas the post-2010 decline dominates the shorter-period analysis.”

We slightly revised the criteria for streamflow data availability used in Figures 2 and 3. In the previous manuscript version, monthly averages were calculated for months with at least 50% data availability, followed by computation of water-year averages. To ensure consistency with other parts of the manuscript, we now apply a uniform criterion: a given year is excluded if more than 10% of the data are missing. Additionally, in Figure 3 (AO and NAO correlations), we replaced the previous restriction to gauges starting before 1980 with a simpler requirement of at least 30 valid years of data. This adjustment allowed the inclusion of additional gauges in the correlation analysis.

- 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.

We agree that the analysis of AO/NAO influence on streamflow variability should be more clearly positioned as part of the study objectives. We have revised the final paragraph of the Introduction accordingly. Specifically, we now state that the study revisits previously reported correlations between Icelandic streamflow and large-scale atmospheric circulation (AO/NAO) using a broader and more recent observational dataset, and we explicitly include this as one of the three key research questions.

Further, we have moved the sentence “The AO and NAO indices were obtained from the NOAA Climate Prediction Center (NOAA CPC, 2024).” from Sect. 3.1.3 to Sect. 2.3 (Data).

- Section 3.1.3: What is the common factor for catchments showing high correlations?

A common factor is that the catchments are not glaciated. We have added the following text to explain this: “Rivers draining the three largest ice caps exhibit negative or near-zero correlations with both the AO and NAO. This pattern reflects that positive phases of these circulation modes

enhance winter snowfall accumulation on glaciers, thereby reducing summer meltwater contributions to streamflow.”

- 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.

We have clarified this in the first paragraph of Sect. 3.2 and removed repeated information about the source of the meteorological data.

- Figure 6: Are these trends calculated based on the grid averages for the whole of Iceland?

Yes, we have clarified this in the figure caption.

- Figure 6: Make sure to focus the presentation/discussion on significant trends, emphasising that the warming signal is more pronounced in the longer period.

We have fixed this in the text.

- 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.

We have added heatmaps for the meteorological trends into the supplement.

- 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.

We thank the reviewer for pointing out the need for clarification. Our intention was not to suggest that higher temperature or ET reverses the spring streamflow trend, but rather that it moderates its magnitude. While most catchments show positive spring streamflow trends, the size of the increase is smaller in basins with stronger warming and ET trends, leading to negative correlations. In other words, warming and higher evaporative demand partially offset increasing spring flows, rather than causing declines. To clarify this point, we have revised the paragraph as follows:

“In the spring season, 20 out of 25 gauges show positive streamflow trends... However, the magnitude of these increases is smaller in basins experiencing stronger warming and higher evapotranspiration trends...”

- Higher temperature may both lead to an increase (snow/glacier melt) and a decrease in streamflow (higher ET). The discussion should better balance this.

We agree, see a new paragraph at the end of Sect. 4.2.2.1

- Figure 9: it is not commented on which of these trends are significant.

We have placed an updated version of this figure in the Supplement, now using glacier areal changes in percentage rather than in km².

- 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?

We thank the reviewer for pointing out the need to clarify the seasonally contrasting role of the baseflow index. To address this, we have added a dedicated subsection (Section 4.3.3) explaining how baseflow contribution influences streamflow trends differently depending on season. In summary, groundwater-fed rivers (high baseflow index) buffer against summer drying, leading to weaker declines, whereas surface-fed rivers respond more directly to rainfall and snowmelt inputs in winter and spring, resulting in larger proportional increases. This seasonal dependence is now clearly stated in the manuscript.

- You may consider reducing the use of ‘in this study’ in the text.

Thank you for pointing out the overuse of “in this study” in the text, we have removed most instances.

- The correlation heatmaps in the Supplement lack reference to significance.

We had specified this in the main text, the final sentence of Sect. 2.6 reads: “Only statistically significant correlations ($p < 0.05$) are reported.”. We have now also made this clear in the Supplement itself (in the figure captions).

- 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

We have added correlation matrices for the catchment attributes (Figures S14 to S18).

References

Rögnvaldsson, Ó., Chang, H., Stanislawska, K., and Okrzesa, M.: Comparison of observed and simulated weather - Description of three dynamical downscaling experiments and comparison with observations, Technical report, available at <http://ftp.belgungur.is/publications/ObsSim-comparison-TechReport2025.pdf>, 2025.