Reply to the Reviewer #1 comments for the manuscript entitled "Topography significantly influencing low flows in snow-dominated watersheds"

The manuscript "Topography significantly influencing low flows in snow-dominated watersheds" from Li et al, submitted to Hydrology and Earth System Sciences, attempts to assess the role of topography on various flow variables to identify important topographic indices (TIs). This was done by first defining a set of TIs of which redundant ones were excluded based on a factor analysis. The remaining TIs were used to estimate their relative contribution to the flow variables: each variable separately for each year. This contribution was estimated with linear regression models. This analysis indicated that the contribution of the TIs were greater for the low flows; the most significant TIs were: perimeter, surface area, openness, and terrain characterization index and slope length factor. Hence, the topography significantly influences the low flows for this study region.

The topic addressed in this manuscript is very interesting. In general, the paper is well written: there is a clear structure, complete and concise summary and a clear title indicating the conclusion of this paper. The methods are generally explained clearly with just a few missing details needed for reproducibility, for example on occurrence/contribution of TIs and on the linear regression model. Based on the results shown, it is clear how the conclusions are taken. However, it is not possible to verify everything, for example the selection of non-redundant TIs. There are just some drawbacks that should be improved.

Response: Thanks for your comments on our manuscript.

General comments:

1) The focus seems to be on having an extensive list of TIs rather than flow variables while both is necessary to thoroughly assess the influence of topography on the flow. Additional flow variables that might be interesting to include are, for example slope of flow duration curve, rising limb density, auto-correlation function or the timing (start of a season, duration of a season).

Response: Thanks for your suggestion. The main theme of this manuscript is to investigate how topography influences several key variables in association with flow regimes. We agree that the suggested additional hydrological variables (slope of flow duration curve, rising limb density, auto-correlation function or the timing, etc.) are also important and should be assessed, but they will be considered in our future studies.

2) (a) Good idea to exclude redundant signatures, but as a result some signatures are only included indirectly (line 260-262) making it rather difficult to assess their significance on the flow variables. (b) Also, what if a secondary TI selected for the analyses (e.g. DDG), turns out to be significant only because of the primary TI (e.g. slope) included in this secondary TI? So how conclusive are the results with the selection of these TIs?

Response: We think both questions are valid and relevant, but they are all related to the test methods themselves we used in this study. (a) The Factor Analysis (FA) is a widely used technique to reduce larger datasets that consist of several variables into fewer sets or factors (Yong and Pearce, 2013). The goal of the FA in this study is to determine the reduced number of significant variables to represent the large dataset of TIs. In our study, the number of

topographic indices (TIs) was reduced from 22 to 11 by the FA. 11 TIs were further classified into two groups and named as "area" and "complexity". Some primary TIs may be indirectly included in the secondary TIs, which indicates that those primary TIs are not significant, and less representative than those secondary TIs. As such, the contributions of these primary TIs not selected by the FA test are much lower than the selected secondary TIs.

(b) Of the selected TIs, both primary and secondary TIs were determined by the FA test. The FA test can distinguish the variables according to the properties of TIs rather than the categories of the TIs (i.e. primary or secondary). In this study, we set three criteria (see the manuscript for the details) to exclude the redundant TIs. We selected the more conservative KMO and anti-image correlation values of 0.7 instead of the commonly-used 0.5. Such selection ensures our results more robust. In this way, the TIs containing redundant information were excluded, while the most representative and significant TIs were kept. Thus, it is possible that some variables (e.g., slope) might be excluded because they are not selected by the FA test, but their signatures may be included in other selected and significant TIs.

3) (a) Explain the methodology of the stepwise linear regression model detailed: include whether a fixed order for including additional TIs was used and whether this order is of significance. (b) In the linear regression, TIs were included even though their resulting estimates are very low. Aren't these TIs insignificant and shouldn't they therefore be excluded?

Response: (a) Thanks for your suggestion. We will include a more detailed explanation regarding multiple linear regression models (MLR) in the revised manuscript. Here, we brief the procedure of MLR to response the reviewer's comments. All the selected 11 TIs by the FA test were initially included in the MLR model. The ANOVA test was used to identify the significance between TIs and flow variables. If one TI was insignificant, then it was removed from the model. The ANOVA test was then re-run for the rest TIs to ensure that all significant TIs were selected. By the trial and error process, final models with only significant TIs to hydrological variables were determined.

(b) The reviewer noticed that some TIs in regression models were not statistical significant (P>0.05), but still included. We agree with the reviewer's assessment, and will exclude those insignificant TIs and update our results accordingly in our revised manuscript.

4) Suggested addition to the discussion: under what conditions are the results transferable to other watersheds? For example are results expected to be differently for even larger watersheds (e.g. 150 000 km^2) with similar climatic conditions?

Response: Thanks for your suggestion. In this study, watersheds sizes range from 2.6 to 1780 km^2 under a similar climate condition. Therefore, our results can definitely be transferred to any other watersheds with the similar watershed size range defined in this study. It is an interesting and open question if our results can be extrapolated to other very large watersheds. We believe that further research is needed to answer this question.

Specific comments:

5) Line 42: The term “somehow” is not nice in a paper, it just be removed here

Response: We will exclude it in our revision.

6) Line 94: include an evaluation with ground measurements if possible

Response: ClimateBC model has been validated by climate stations across British Columbia, Canada. The validation process is presented in Wang et al. (2006), and this citation is cited in the manuscript.

7) Line 119: unclear unit of the flow variables: mm/year or mm/d?

Response: The unit is mm/d. We will correct it in the revised manuscript.

8) Line 129/130: include in section “2.1 Study Watersheds” that the topography is similar and how that is assessed.

Response: We assumed that topography is similar if the average values of TIs at the watershed level do not vary significantly. This is because that the average values can cancel out large variations in TIs. As such, we used standard deviation to represent the variations in TIs between watersheds.

9) Line 134: The third criteria seems to be excluded from the analysis. It is not mentioned after this section. If that is the case, then exclude it here too.

Response: No. Excluding a TI by the FA test is a trial and error process. Therefore, we strictly followed all three criteria. Therefore, this criterion was not excluded. We will make it more explicit in the revised manuscript.

10) Line 158: It is unclear what exactly is meant with the “occurrence of a TI” and how it is determined (it’s actually clearer in the caption of Fig 3).

Response: Original Line 158: *O* is the number of occurrences of a TI in a flow variables model.

Revision: *O* is the number of selected TIs appearing in the final multiple linear regression models.

11) Line 160: It is unclear what exactly is meant with the “contribution of each TI” and how it is determined.

Response: The “contribution of each TI” means the contribution (%) of each selected TI to explain the R^2 of each multiple regression model.

We have provided such statement in Line 153-155. The R package “relaimpo” was used to quantify the relative contributions of the selected TIs to flow variables. Specifically, the relative contribution of each independent variable to R^2 was calculated for each model.

12) Line 167: Include results indicating that other (groups of) TIs are indeed repetitive and should be excluded for the sake of verification.

Response: We understand your comment. Although 11 TIs were excluded by the FA test, it would be useful to indicate them. Thus, we will highlight the selected TIs in Table 1 so that those unselected TIs can easily be identified.

13) Line 170: a) Variance of what? b) How is it calculated?

Response: a)

Revision: First and second factors explained 80.9% and 11.7% of <u>total variance of TIs in the selected watersheds, respectively</u>.

b) The total variance is one of the standard outputs of the FA. In this study, it is calculated by the SPSS software.

14) Line 170-175 and Fig2: SA is the surface area, yet it is not grouped in group 2 which describes the area!

Response: As described in Table 1, the SA is the land area of each DEM pixel. It represents the roughness of a DEM pixel or a watershed. The larger value of SA indicates more roughness of a watershed. Therefore, it is classified in Group 2.

15) Line 182: Confusing formulation of "1, 1, 4, 8, 9, 10 and 11"

Response: The Kendall tau tests were conducted between flow variables and each TI for each year from 1989 to 1996. Due to the difference in climate among years, flow variables showed distinct responses to topography. Therefore, we counted the numbers of the significant TIs to represent the relationship between flow variables and TIs. Here is the revision.

<u>Revision: The number of significant TIs in each year increased from 1 to 11 from 1989 to 1996.</u>

16) Line 194: Sentence contradicts results in Fig 3. According to Fig 3, SA does not play a significant role in Q_{min}, but is significant in $Q_{75\%}$.

Response: Thanks for pointing this out. We have corrected this statement.

<u>Revision: SA played a prominent role in $Q_{75\%}$, but did not significantly contribute to the variation of Q_{min}.</u>

17) Line 221: "positive relationship between the selected TIs and low flow variables"; this is not the case for the openness which always has negative estimates in the regression model results as shown in the supplements.

Response: Thanks for pointing this out. We will revise our statement accordingly in the revised manuscript.

18) Line 256: It is not surprising these commonly used TIs such as slope was not in the final list as it was excluded from the selected TIs and therefore its contribution was not even calculated in order to end up in the final list. Its contribution is only indirectly assessed through other TIs. How different would the results and conclusion be if all TIs were included?

Response: To answer this question, we have re-run models for the minimum flows for three years without conducting the FA test first. The other procedures were kept the same. The Table R1 in this response letter (below) showed that the models now included several redundant TIs (e.g., wetland coverage, roundness, stream length) if the FA test was not conducted. However, they were not significantly correlated with low flows indicated by Kendall correlation test. This further suggests that some redundant and nonsignificant TIs

would be included if the FA test was not conducted initially. It further proves the benefits of adopting the FA test.

Table R1. Topographic indices included in the multiple regression models of Q_{min} for years of 1989, 1990, and 1991 with or without the FA test.

Year	Model variables with initial FA test	Model variables without initial FA test
1989	**LS,** Openness, **Perimeter**, **Slope**, TCI, **TRI**, UCA	DDG, DDGD, **LS**, Openness, **Perimeter**, Relief, Roundness, SA, Stream Length, **Slope**, **TRI**, TPI, Wetland
1990	**LS, Perimeter**, Relief, SCA, Slope, **TCI**, **TRI**	Median Elevation, Roundness, Stream Length, Wetland, TPI, **LS**, Openness, SA, **TCI**, **TRI**
1991	**LS, Perimeter**, **Openness**, Slope, TCI, UCA	DDG, DDGD, LS, **Openness**, **Perimeter**, Relief, Roundness, Stream Length, SCA, SA, TRI, Wetland

Note: the TIs (in bold style) are the common TIs between two models.

19) Fig 1: hydrometric stations are plotted, yet not mentioned in the paper. It is suggested to either mention how they were included or exclude them from the figure.

Response: We have included the stations in the revised manuscript. Please also see response NO. 22.

Technical corrections:

20) Line 170 and 175: Fig 3 written, yet probably referred to Fig 2

Response: Yes. The correct reference should be Fig 2.

21) Line 193: Fig 4-7 missing

Response: This is a typo. We meant Fig. 3 (A-D).

22) Fig 1: missing scale bar for the small map of the state. This small map misses the background map showing the location of neighbouring land, now it seems the state is an island which is not the case!

Response: Thanks for your suggestion. We have redrawn Figure 1.

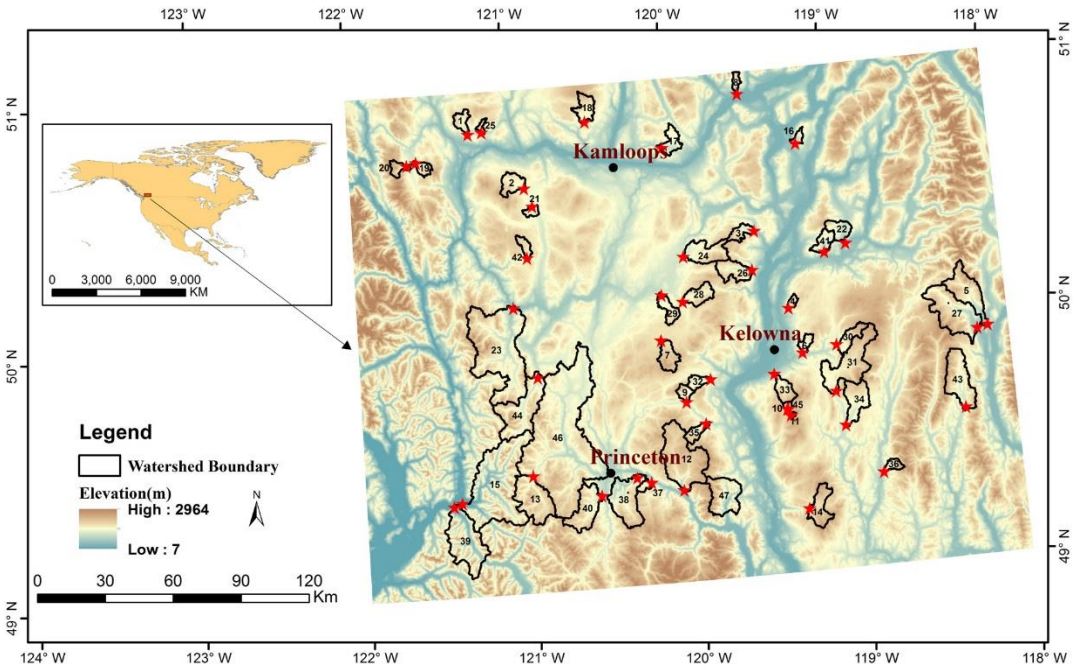

Figure 1. Location, elevation, watershed boundary, and hydrometric station of selected 28 watersheds in the Southern Interior of British Columbia, Canada.

23) Fig 2: missing description of axes (what do these numbers on the axes indicate?); confusing choice of words: factor 2 = group 2. Use the same thing in the figure and label.

Response: First and second factors are the standard expressions of the factor analysis. Based on the factor analysis, we classified the TIs into two groups. To avoid the confusion, we named two groups to "Group A" and "Group B", respectively.

24) Supplement tables list: inconsistent font types

Response: Thanks for pointing this out. We will revise them accordingly.

25) Supplement Fig S1: inconsistent abbreviation for precipitation (PPT in figure and P in capture); do not connect the points to lines as the results for each watershed are independent from each other; line for temperature is not visible in a black/white print

Response: We will revise them accordingly.

Reference

Yong, A. G., Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in quantitative methods for psychology*, *9*(2), 79-94.