Author response to reviewer #1 comments for HESS manuscript "Modelling the effects of climate and landcover change on the hydrologic regime of a snowmelt-dominated montane catchment" [Paper #: hess-2024-361]

Dear Reviewer,

We would like to thank you for your thorough and thoughtful review of our manuscript submission. You raised many important issues that will certainly result in a stronger manuscript. Please find below a list of responses to your comments. We hope our responses satisfy the spirit and intent of your remarks.

Sincerely,

Russell Smith

Reviewer #1 comments

General Comments

This paper combines a large amount of data and different climate and land cover scenarios in a modelling study to determine the combined effects of land cover change and climate change on the snowpack and streamflow regime for a headwater catchment in British Columbia, Canada. This is a huge undertaking and I appreciate that the effects are analysed for many different aspects of the snowpack and hydrograph. The paper clearly shows the interaction between the effects of land cover change and climate change. For some aspects of the hydrograph or snowpack, the land cover change enhances the effects caused by climate change and for others, it mitigates it. The results furthermore highlight the importance of the location of the disturbance in the catchment (i.e., whether the vegetation is replaced in the upper or lower part of the catchment) and the time since the disturbance. These are important results and highlight the need to consider the effects of land cover and climate change jointly and to not study the effects of land cover change for only one climatic period.

- Thank you for the very positive comments, and for acknowledging the large effort.

Unfortunately, some of the model decisions are not so clearly described and it is not very clear how the model was calibrated. There is also no mention of the uncertainties in the results due to parameter uncertainty. Considering the potentially very large number of parameters that are optimized, it is possible that a different parameter set would lead to considerably different results. The lack of uncertainty analyses is acknowledged in the final part of the discussion, but I would argue that at least some model uncertainties need to be presented. As a result of the lack of a clear description of the calibration procedure and the lack of an uncertainty analysis, it is not clear how the presented results are influenced by the model decisions or model parameter sets (equifinality).

- In your comments above and below, you point out several areas in the manuscript where the optimization process could be clarified. We acknowledge the need to provide clarity around the issues you identified, and have addressed your related comments in the sections below. We propose increasing our discussion of uncertainties in the manuscript (Section 6.4) to address parameter uncertainty in greater detail. We will also provide more details in the manuscript from the points below related to the calibration process.
- We further acknowledge the importance of considering uncertainties, including variability in meteorology and landcover (which we addressed thoroughly), and parameter uncertainty. We acknowledge that a comprehensive parameter uncertainty analysis would be an asset to the manuscript; however, it would be a very large undertaking, on top of the very large undertaking we have already completed (as you acknowledged). While your request for a parameter uncertainty analysis may seem like a modest undertaking, it is important to consider the amount of modelling and data synthesis involved. In this respect, consider the multiplication effect of eleven climates, five landcovers, several key hydrologic variables (snowpack accumulation, snowmelt timing, peak flow, annual yield, low flow, extreme events), numerous alternative parameter sets, and 100 years of daily simulation for each combination of these. Additionally, we calibrated the model using the Dynamically Dimensioned Search algorithm (Section 2.3.4 in manuscript). To generate a selection of alternative parameter sets for the uncertainty analysis, we would need to recalibrate the model using a randomized search algorithm (e.g., Monte Carlo simulation). As noted below, we believe the model was constrained well by the calibration process, resulting in limited parameter uncertainty. Furthermore, we do not have the resources (funding or time) available to complete such a parameter uncertainty analysis. It is also important consider that the paper is already long, as you acknowledged. We found it challenging to clearly and effectively present the results that are already provided. Accounting for uncertainty (e.g., error bars, additional lines or plots) would exacerbate that challenge.

- Further to the points above regarding parameter uncertainty, we believe the model was constrained well by the calibration process, resulting in much less parameter uncertainty than is suggested. All model parameters were calibrated simultaneously on the SWE and discharge datasets shown in Figures S2.1 through S2.5 (manuscript supplement) and on annual precipitation at Penticton Airport, then validated on the datasets shown in Figures S3.1 through S3.5. It is a rich optimization dataset that encompasses a large range of:
    - intra-seasonal and inter-seasonal meteorological and hydrological conditions (e.g., 2010 vs. 2011 SWE; 1972 vs. 1973 discharge in Penticton Creek; 1986 vs. 1990 vs. 1992 discharge in 240, 241, and Dennis Creeks, see Figures S2.1 through S2.5),
    - catchment forest cover conditions (e.g., 240 Creek Sub-catchment vs. 241 and Dennis Creek Sub-catchments, see Figures 1 and 2 in main body),
    - catchment elevations and scales (e.g., Penticton Creek Watershed vs. sub-catchments),
    - catchment orientations (e.g., 241 Creek vs. Dennis Creek Sub-catchments),
    - forest cover conditions at SWE stations (e.g., UP13 vs. UP2, UP9 vs. UP10, see Figure S2.5; UP1 and UP3 vs. UP2 and UP4, see Figure S3.5),
    - elevations of SWE stations (e.g., UP13 vs. UP9, 2F08 vs. UP10), and
    - orientations of SWE stations (e.g., UP9 vs. UP13).
- To support the parameterization, a forest cover survey was completed that encompassed seven mature forest plots ranging in elevation from 649 m (dry, low elevation forest) to 1930 m (wet, sub-alpine forest) (locations provided in Figure 1). The sampling protocol was based on Canada's National Forest Inventory Ground Sampling Guidelines (Canadian Forest Inventory Committee, 2008). Measured variables included leaf area index, crown closure, tree diameter, tree height, and species, among others. Leaf area index and crown closure (particularly important in the model) were measured at 19 points per plot using hemispherical photos.
- A composite objective function was utilized (see Table S3 in supplement) that allowed the optimization to be constrained well by focusing on different features of the optimization dataset. Each data feature and component of the objective function were important for different reasons, and provided value for constraining different model parameters. For instance:
    - Overall yield (i.e., absolute bias) and variance (i.e., NSE) were applied separately, to ensure that neither could be fit well at the expense of the other.

  - Absolute bias constrained parameters controlling overall water volume (e.g., precipitation lapse rates, evapotranspiration in winter and summer).
  - NSE constrained parameters controlling the timing of snowmelt and runoff (e.g., energy balance, runoff routing).
- To ensure that internal model processes were functioning well, individual components of the composite objective function were applied separately to discharge and SWE, smoothed and unsmoothed discharge, spring freshet and low flow, and sub-catchments and the main catchment outlet.
- The table below outlines factors that were important for constraining key parameters. The constraining features identified in the table relate to features in the optimization dataset that were particularly important for constraining the corresponding parameter(s).

<table>
<tr><th>Parameter type</th><th>Parameter or parameter group</th><th>Constraint</th></tr>
<tr><td rowspan="2">Short-wave radiation</td><td>CloudTempRanges</td><td>Calibration range was based on a comparison of measured short-wave radiation and diurnal air temperature for P1 climate station.
Constrained by calibration on SWE and discharge.
Constraining features:
• rate of snowmelt and associated runoff volume during intra-seasonal sunny periods versus cloudy periods
• snowmelt timing and associated runoff timing in wet vs. sunny spring freshet seasons
• differences in snowmelt and runoff between different slope aspects / catchment orientations.</td></tr>
<tr><td>UBCCloudPenetration</td><td>Calibration range was based on ratio of low mid-day short-wave radiation vs. high mid-day short-wave radiation for P1 climate station.
Constrained by calibration on SWE and discharge.
Constraining features:
• rate of snowmelt and associated runoff volume during sunny periods
• differences in snowmelt and runoff between different slope aspects / catchment orientations.</td></tr>
</table>

Air temperature	AdiabaticLapseRate, WetAdiabaticLapse, UBCTempLapseRates	Calibration range was based on lapse rates calculated using weather data from P1 and Penticton Airport stations. Constrained by calibration on SWE and discharge. Constraining features: • differences in snowpack accumulation and snowmelt rates at different elevations during periods when air temperatures are close to zero degrees • differences in snowpack accumulation and snowmelt rates at different elevations for wet vs. dry periods • differences in volume and timing of runoff during snowmelt periods for sub-catchments vs. main catchment outlet
	ReferenceMaxTemperatureRange	Calibration range was based on diurnal air temperature for P1 climate station during clear sky conditions, and based on values suggested in the user manual for UBC Watershed Model (Quick, 1995). Constrained by calibration on SWE and discharge. Constraining features: • rate of snowmelt and associated runoff volume during intra-seasonal sunny periods versus cloudy periods • snowmelt timing and associated runoff timing in generally wet vs. sunny spring freshet seasons
Precipitation	PrecipitationLapseRate	Calibration range was based on calculation of long-term precipitation lapse rate between Penticton Airport and P1 climate stations. Constrained by calibration on long-term mean precipitation at Penticton Airport. Constraining features: • rate of snowpack accumulation in clearings during snowfall events • volume of runoff during rainfall events
	RainSnowTransition	Values assigned based on those suggested in the user manual for UBC Watershed Model, and based on values used in BC Hydro operational models.

Snowpack albedo	UBCSnowParams (P0ALBMIN, P0ALBMAX, P0ALBREC, P0ALBASE, P0ALBSNW, P0ALBMLX)	Calibration ranges based on values in literature (e.g., Spittlehouse and Winkler, 2004) and values used in BC Hydro models. Constrained by calibration on SWE and discharge. Constraining features: • differences in rate of snowmelt and associated runoff volume during sunny periods after recent snowfall vs. after several days without snowfall
Maximum liquid water content of snow	IrreducibleSnowSaturation	Value assigned based on values in user manuals for Raven (Craig and the Raven Development Team, 2022) and UBC Watershed Model, as well as values used in BC Hydro models.
Snowpack cold content	CC_DECAY_COEFF	Calibration range was based on values suggested in UBC Watershed Model and values used in BC Hydro models. Constrained by calibration on SWE and discharge. Constraining features: • rate of snowmelt and associated runoff volume and timing after periods of cold weather
Snowpack patchiness	SNOW_PATCH_LIMIT	Calibration range was based on visual observations of snowpack patchiness and snowpack survey data. Constrained by calibration on SWE and discharge. Constraining features: • spatially averaged rate of snowmelt and associated runoff volume when snowpack is nearing complete melting
Soil layers and thickness	SoilProfiles	Calibration ranges based on soil mapping for the catchment. Constrained by calibration on discharge. Constraining features: • Runoff volume and timing, particularly in comparing sub-catchments with different soil distributions, and in comparing spring freshet runoff vs. rainfall runoff during snow-free periods • Simulated evapotranspiration and, thus, low flow volume are sensitive to soil depth

Soil texture and porosity	%SAND, %CLAY, %SILT, %ORGANIC, POROSITY	Values assigned based on soil mapping for the catchment.
Soils / runoff routing (infiltration/runoff partitioning, percolation, interflow, baseflow)	HBV_BETA, MAX_PERC_RATE, MAX_INTERFLOW_RATE, BASEFLOW_COEFF, BASEFLOW_N	Calibration ranges based values suggested in Raven user manual and input from Raven developers (J. Craig, personal communication, September 28, 2018). Constrained by calibration on discharge. Constraining features: • Runoff volume and timing (e.g., flashiness), particularly in comparing sub-catchments with different soil distributions, comparing sub-catchments vs. main catchment outlet, and comparing spring freshet runoff vs. rainfall runoff during snow-free periods • Shape of spring freshet hydrograph and rainfall driven event hydrographs • Rate of runoff recession at different points in time after peak flow • Low flow volume
Potential evapotranspiration	PET_CORRECTION, PET_VEG_CORR	Constrained by calibration on SWE and discharge. Constraining features: • snowpack accumulation in forests vs. clearings • spring freshet and low flow runoff volume in 240 Creek Sub-catchment (mature forest dominated) vs. 241 Creek and Dennis Creek Sub-catchments (extensive forest cover disturbance), and in sub-catchments vs. main catchment outlet.
Tree height	MAX_HT	Values assigned based on forest cover mapping. Model is insensitive to tree height.
Crown closure	FOREST_COV	Values assigned based on vegetation surveys and forest cover mapping.

Leaf area index, forest shading	MAX_LAI, SVF_EXTINCTION	Calibrated for dominant forest types. Assigned for minor forest types based on calibrated results for dominant types. Model is generally insensitive to the minor forest types. Calibration ranges based on vegetation surveys, forest cover mapping, and values in literature. Constrained by calibration on SWE and discharge. Constraining features: • snowpack accumulation, snowmelt rate, and snowmelt timing in forests vs. clearings • timing of hydrograph rising limb during spring freshet in 240 Creek Sub-catchment (mature forest dominated) vs. 241 Creek and Dennis Creek Sub-catchments (extensive forest cover disturbance), and in sub-catchments vs. main catchment outlet
Canopy interception	TFRAIN, TFSNOW, MAX_CAPACITY, MAX_SNOW_CAPACITY, MAX_INTERCEPT_RATE	Calibrated for dominant forest types. Assigned for minor forest types based on calibrated results for dominant types. Model is generally insensitive to the minor forest types. Calibration ranges based on vegetation surveys, forest cover mapping, and values in literature. Constrained by calibration on SWE and discharge. Constraining features: • snowpack accumulation in forests vs. clearings • volume of spring freshet runoff and volume of runoff during rainfall driven events (e.g., rain-on-snow) in 240 Creek Sub-catchment (mature forest dominated) vs. 241 Creek and Dennis Creek Sub-catchments (extensive forest cover disturbance), and in sub-catchments vs. main catchment outlet
Leaf conductance	MAX_LEAF_COND	Values assigned based on those suggested in Raven user manual, and based on values used in BC Hydro models.
Terrain attributes	HILLSLOPE_LEN, DRAINAGE_DENS	Values assigned based on field observations, stream mapping, and values used in BC Hydro models. Model is insensitive to these parameters.

Stream channel geometry and grade	SurveyPoints, Bedslope	Values assigned based on field observations and digitizing in Google Earth
Stream channel roughness	RoughnessZones	Initial values based on standard Manning's roughness coefficients, then allowed to vary in the calibration to match the travel time of flow between the sub-catchments and the main catchment outlet.

The graphs used to present the results are clear and very useful. But it would be good if they had error bars to represent the range of results caused by equally good fitting model parameter sets.

- See relevant comments above and below.

I like it that the time series of the simulated and observed runoff are given for the individual years in the Supplementary material.

- Thank you for acknowledging the value of the time series. We would like to point out that many papers are published with much less detail related to the simulated and observed time series (e.g., fewer and/or much smaller plots are often provided). We provided this detail so the reader could clearly review the fit to the observed data, as the observed data were very important for constraining internal model processes. We believe that providing these time series for review adds credibility to the results.

The paper is long but overall, well written.

- Thank you for this comment. We agree that the paper is long. Because of its length and the large amount of results/data, we've had to put a large effort into refining the manuscript. This issue relates to some of our concern with adding content to the plots and text to address parameter uncertainty, as discussed above.

Specific comments

L14 and 862: Quantify this in a different way, e.g., in days or weeks. 2-9 times more is important if we talk about an advance of a week or several weeks due to disturbance but not if the advance is only 1 day.

- Good point. We propose revising the text to the following: "The combination of climate change and stand replacing landcover disturbance in the middle and high elevations is predicted to advance the timing of the peak flow two to nine times (depending on emission pathway) more than the advance generated by disturbance alone (7 days)."

L26: Maybe use a different word than values (hydrograph characteristics, hydrological signatures?)

- We are referring to values of concern to society, as changes in hydrology cascade into changes to watershed risk. We will work on clarifying the wording in the manuscript.

L76: Considering all the uncertainties in these assessments, the decimals are probably not warranted here.

- Good point. We'll remove the decimals.

L92: Give some info on the model here already. It would be good to know for the reader early on if you are using a physically based, spatially distributed model or some other model, if it was calibrated or not, etc.

- We'll add a couple points, in line with your suggestion.

L126-129: It is nice that you describe the vegetation here and give the codes that you will use for the vegetation codes throughout the text but it is hard for the reader to remember these codes, especially since there are also codes for the different scenarios. In other words, it would be a lot easier for the reader to understand the parts about the vegetation if you would just write out the names instead of using the codes.

- Okay. We'll go with your suggestion, provided the language does not become too cumbersome in the text.

L139: In addition to the mean annual runoff, also mention the mean annual precipitation, either averaged over the catchment or for at least one station. This is important information about the study site.

- Good suggestion. Thank you.

Section 2.2.1: It would be good to already mention how many HRUs there are in this section (now it is only mentioned on L243) and how many parameters there are per HRU. Now this section is short and a lack of knowledge on the model and its parameters early on in the paper, hampers the understanding of the other parts in section 2.2.

- Thanks for the suggestion. We think that it's best to discuss these details in the "spatial discretization" section, and would like to avoid redundancy. However, we will consider reordering Section 2.2 to have the meteorology section after the landcover and spatial discretization sections.

L195: How well is well? Is there a reference here or a result that you can add to the supp materials?

- We will provide a citation for Spittlehouse and Dymond (2022).

L267: How many parameters are there per HRU and in total? and how many of these were calibrated? Even after reading the paper, this is unclear to me. Please add this information clearly in the methods section. Ideally already in section 2.2.1.

- Each HRU had:
    - 13 soil / sub-basin runoff related parameters (8 calibrated)
    - 13 vegetation related parameters (6 calibrated for dominant forest types; all 13 assigned for minor forest types based on calibrated results for dominant types)
- There were also 30 parameters related to meteorology and energy balance (18 calibrated), and 12 parameters related to in-channel runoff routing (9 calibrated).
- See additional details provided in the table above.
- We will consider opportunities to provide this information concisely in the manuscript.

L268: What weighting did you use for the calibration? Equal for each of these objective functions?

- All components of the objective function shown in Table S3 (supplement) were assigned an equal weighting. This point will be added to footnote #1 for Table S3, and added to the text in the main body.

L281, 283 and ff: What exactly do you mean by constrained (or in L301 and 306 by informed)? Did you pick a parameter value a priori and not calibrate it or did you select a parameter range and calibrate within this range?

- Please see the table above for additional details on setting calibration ranges and calibrating on empirical data.
- The word "informed" was utilized to indicate that the data were used for setting calibration ranges (dominant forest types) and for assigning parameter values (minor forest types). We will revise the text to clarify these points.

L313: This wording is not clear. Did you use it to guess a specific value and then use this in the model? Did you calibrate within a certain range? A bit more information, or clearer wording would be useful.

- Field knowledge was used for adjusting calibration ranges. We will clarify this point.

L321: This is not clear - how did you get values for each specific channel? How different were these values?

- Google Earth was used for digitizing the width of the lower mainstem channel (10-15 m). Visual field observations were used for assigning the width of smaller channels in the upper reaches that were not clearly visible in satellite imagery (2-3 m wide). These points will be clarified in the text.

L331: How many parameters were optimized and how many were fixed based on field knowledge? Also did you use the same parameters for all the HRUs with the same vegetation or soil? Would it be possible to add a table with all parameter values and the range used for the optimization somewhere?

- See details above related to optimizing and assigning parameter values.
- All HRUs with the same vegetation type had the same parameter values for vegetation, and all HRUs with the same soil type had the same parameter values for soils. In this respect, two HRUs could have the same vegetation parameter values, but different soil parameter values, and vice versa.
- With respect to providing a table of parameter values and calibration ranges, we acknowledge the value in being able to review the parameter values. However, the primary author is a business owner in a competitive consulting environment, and catchment modelling of different climate and landcover scenarios forms an important component of his business. There is a large investment of intellectual property involved with parameterizing the model. There would be a considerable risk to his business competitiveness by publishing the model parameters.

L333, 389: How did you weigh these different objective functions in the calibration? All equal weight? Or did you optimize each function individually first? From L323-326, it appears that you did it sequentially? Or did you just use different time periods for each of these objective functions and calibrate everything at the same time using some weighted function? The current description doesn't make the calibration process very clear to me. Also, what is the reason for not using the NSE for the entire study period as well?

- As discussed above, all components of the objective function shown in Table S3 (supplement) were assigned an equal weighting, and all parameters were calibrated simultaneously. The language in L323 was intended to convey that different time periods were used for different data types (e.g., 1971-1981 for Penticton Creek discharge, 1984-1992 for discharge in the sub-catchments, and 1995-1997 and 2009-2014 for SWE). These varying periods of record were related to the availability of data in different time periods. These points will be clarified in the text.
- We should point out that the "clearcut" labels for UP9, UP11, and UP13 in Figures S2.5 and S3.5 should be labeled as "regen" (i.e., regenerating), consistent with Table S2. Also note that "leading species" in Table S2 should be changed to "vegetation type".
- Good question about NSE. It was not used for the entire study period because there is lower certainty in the quality of measured winter discharge related to potential ice build-up on the weir crest. Moreover, for constraining the model, the overall volume of runoff during the low flow period was more important than any short-term minor changes in flow, as the overall low flow volume relates to evapotranspiration and slow soil drainage / runoff processes. For these reasons, the decision was made to focus the low flow calibration on overall yield (i.e., absolute bias). These points will be clarified in the text.

L461: Already mention here if this is largely due to a change in precipitation or due to a change in evapotranspiration.

- Increasing evapotranspiration was an important cause of the decrease in net precipitation (P-E) in the 2050s, and increasing precipitation in the 2080s generated the partial recovery. This will be clarified in the text.

Figure 9: The shape of the curve changes as well. What is causing this? This requires some discussion.

- Good question. It is assumed that this point relates primarily to Figure 9b. The shape of the curve changes because there is a decrease in the peak flow for frequently occurring peak flows, and an increase for extreme peak flows. These changes are driven by the general decreases in snowpack accumulation (i.e., snowmelt runoff) and net precipitation, coupled with increasing extreme rainfall intensity. These points are made in the existing text; however, we will be sure to make a stronger connection to the shape of the curve.

Section 5.2.4: Make it clearer that this is the annual *average* discharge

- Good point. We will make that change.

L598: A lot of the quickflow probably consists of subsurface stormflow or even groundwater flow. The majority of quickflow is unlikely to be overland flow (surface runoff) for a forested catchment.

- Agreed. The point being made in this line relates to the forested condition generating an increase in event frequency for annual discharge, whereas it’s the large burn that generates an increase in event frequency for peak flows.

L780: Groundwater would be a more likely source for the streamflow in the dry period than soil water (retention).

- Agreed. We will clarify this point in the text.

Minor comments

L11: Mention the name of the model or the type of model in the abstract.

- We'll add that point. Thank you.

L82-89: Move to the study site description.

- We believe this physiographic information provides helpful context upfront that landcover varies considerably in space and time.

L121: Explain that BEC is the biogeoclimatic ecosystem classification.

- Thank you.

L191: Lowest temperatures instead of coolest temperatures.

- We'll make that change. Thank you.

Figure 5: Maybe still add South and North to the axis labels for subpanel b?

- A few reviewers have requested clarification for this figure. We will rearrange the figure and adjust labelling to make it easier to interpret.

L431: These differences are very small. Highlight that first before giving the values!

- Good point. We'll make that change. Thank you.

L700: What do you mean by snowpack loads?

- We mean snowpack accumulation. We'll revise this for consistent wording.

L720-721: Explain better how this sentence fits here / what you mean by this? What is the link to the previous or next sentence?

- This point was provided to relate increasing rainstorm intensity to more rapid hillslope runoff, which would increase peak flows. We'll revise the text to clarify.

L797 values at risk: Do you mean the streamflow signatures / hydrograph characteristics? This could be worded more clearly.

- See related comments above.