

Dear editor,

Thank you for your detailed comments. We have revised most of the figures (there are now less subplots to keep the overall figure count as low as possible), added some additional figures where requested and tried to make this paper more transparent. We added more background information on the calibration process that was used as part of the implementation of WaSiM and clarified the role of permanent snow and ice in the Kawarau catchment. The potential issue of snow towers accumulating along the main divide and the headwaters of the catchment was also investigated. We hope that this revised version of the paper now provides sufficient information on the hydrological model as well as the snow and ice stores during both the reference and the future periods.

Best regards,
Andreas Jobst (on behalf of all co-authors)

Editor Decision: Publish subject to revisions (further review by editor and referees) (10 Mar 2018) by Kerstin Stahl

Comments to the Author:

Dear authors,

thanks for the revisions. I found several of them satisfying, but when reading the paper in detail now, I found there are crucial remaining issues related to the reviewer comments, which I do not see addressed sufficiently in the replies and added material. Also there are number of smaller editorial aspects relating to figures etc. that need to be improved, before I can make a final decision. I will keep the option open to appoint another reviewer again to advise.

Referee 1 in points 2 (*original comment: "Catchment characteristics: Please provide some information about i.e. elevation range, glacier extent, vegetation of the sub-catchment (outlet Chards Rd)"*) and 3 (*original comment: "Snow models: The study focusses on snowmelt uncertainty for hydrological projections and should therefore give more details about the two types of snow models used here. For example, it would be helpful to know if all snow is gone after the summer (except for glaciers) or if snow accumulates from year to year."*) raised questions about detailed land cover and about the potential influence of snow accumulation over the years. Referee 2 found that streamflow is not enveloped in the simulation (*original comment: "The range of simulations in the paper do not envelope the observed flow – so there is some uncertainty that is unaccounted for in the paper and I am left wondering where it is? Some additional discussion is warranted here, including discussion of potential uncertainty in hydrologic model parameters."*) and the reviewers lacked some important background material. I think the comments are very important for the interpretation of the results and are not sufficiently elucidated and not transparently addressed. And they might be related.

The added short description and the new figure with some example time series of SWE are both not sufficient. In particular, it is not proven whether so-called snow towers (e.g. see Freudiger et al. 2017 in WIRES Water for a review) build up only in the accumulation area of glaciers and nowhere else (in the past but in particular also in the future runs). The manuscript needs to be more informative and transparent about spatial and long-term cumulative, quantitative, information about cryosphere water balance terms.

Response A: We have added a new figure (Figure 4 – also included here) that includes a snow tower metric to investigate the build-up of snow across melt seasons (i.e. “snow towers”) in our study. The snow of glacier-free cells does not melt completely over summer and autumn but neither does it show a positive trend which suggests that the residual snow does not keep growing from year to year (plot b). We also added a snow tower statistic (plot a) that shows the combined area of the cells at the end of each hydrological year (31st of march) where SWE is greater than the amount of snowfall during the corresponding hydrological year. Only very few cells show such snow towers but it is interesting to see that the Tindex model produces more snow towers than the Anderson model. The effect on the water balance of these snow towers is rather small (< 2.5% of the entire area is potentially affected) but it would be interesting to use the latest WaSiM version that comes with a snow redistribution routine as part of future work. We have added the following material to the manuscript.

P6 L10 – P6 L21: “Due to the high annual precipitation totals the grid cells in the upper catchment could potentially be prone to large build-ups of snow that persist throughout the melting period and into the next hydrological year. As discussed by Freudiger et al. (2017) such snow towers can cause substantial errors in the modelled water balance and should therefore be made transparent. As indicated in Figure 4b the snow pack of some cells persists across all of the melt seasons with SWE values at the end of the melt season ranging between 7 and 46 mm (21 and 77 mm) for WaSiM-Anderson (WaSiM-Tindex). To further investigate the significance of this potential issue, all cells where the SWE value at the end of the hydrological year exceeded the annual snowfall (excluding glacier cells), were marked. The combined relative area of these cells is shown in Figure 4a and does not exceed 1% for WaSiM-Anderson, while for WaSiM-Tindex the affected area does not exceed 2% except for the years 1997 and 1998. The mean exceedance (SWE/snowfall) for the individual years based on all affected cells varies from 1.3 to 1.8 for WaSiM-Anderson and from 1.4 to 2.3 for WaSiM-Tindex. Thus, for a number of cells along the main divide the modelled build-ups of snow could be classified as potential snow towers. However, the combined area of these cells is small, which means that the effect on the modelled water balance is negligible.”

P11 L32 – P12 L2: “Accounting for snow redistribution by avalanches or wind as described in Freudiger et al. (2017) should also be investigated as this would potentially reduce some of the relatively large accumulations of snow that were found to persist over multiple years on some cells in the upper catchment during the reference period.”

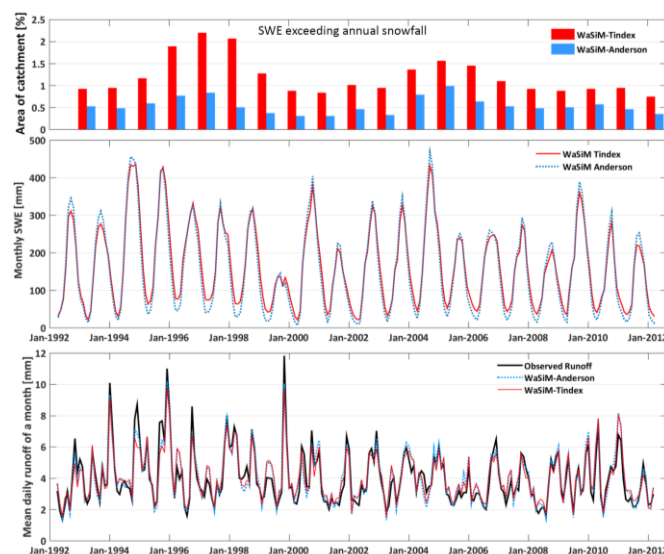


Figure 4 (a) relative area of the Kawarau catchment where SWE exceeds annual snowfall for a hydrological year (e.g. 1.4.1992 – 31.3.1993) as an indicator for unrealistic build-up of snow (i.e. snow towers). (b) Modelled mean monthly snow water equivalent (SWE) based on the Tindex and Anderson simulations. (c) Observed and modelled (WaSiM-Anderson and WaSiM-Tindex) monthly runoff at Chards Rd for the reference period (1992-2012).

We have also added the following figure to the manuscript which shows the mean SWE storage for the six months with historically the greatest snow storage across the 21st century.

P10 L7 – P10 L12: “As shown (Figure 11) by the transient simulations of mean SWE from July to December (months with historically the highest SWE) there is no clear distinction between the A1B and A2 simulations until the last quarter of the 21st century when the median of the A2 runs stays consistently under the A1B runs. While the median SWE based on all 16 QM runs decreases substantially (60%) during the 110-year period, years with large accumulations of SWE still occur in the second half of the 21st century for individual members and years. The negative trend was found to be more pronounced for the glaciers in the catchment, which are projected to lose 93% of their volume (91% based on A1B and 94% based on A2) by 2099.”

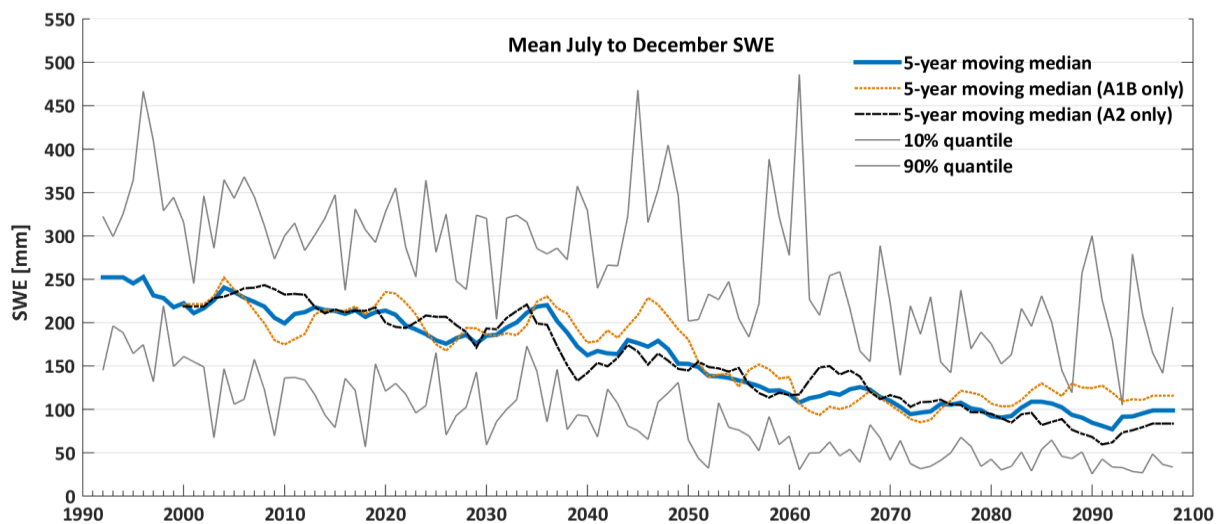


Figure 11 The mean SWE of the six months with the historically highest SWE (July to December) is shown for the 16 transient QM simulations (including QM-Anderson and QM-Tindex).

The main aim is climate change assessment and hence long-term changes. Hence, this long term change needs to be the benchmark for the model simulation in the past. In such an environment long-term accumulation of snow and the correct glacier evolution (such as retreat) are interdependent and will be primary determinants of long-term changes to the water balance (and thus changes in the hydropower potential which is cited as the underlying motivation).

Indeed, now after learning from the revised text that there is considerable glacier area in the modelled catchment, I need to request more modelling details a) how snow redistribution is handled in WaSiM and b) how the glacier(s?) and glacier change is/are treated within the model. Please also add exactly how much the glacierized area is in the here-shown modelled catchment, not the entire Clutha (as the new sentence says now) and in particular how much the streamflow contribution from glacier mass balance changes is. A correct representation of the long-term glacier change will be the best proof for a generally correct snow model as well!

Response B: We have added the following sentences to the manuscript to clarify how much ice volume is estimated to be found inside the Kawarau catchment (P4 L4 – P4 L9): “The ice-covered area inside the Kawarau catchment is small and based on 2001/2002 satellite imagery amounts to 1.8% or 84 km² (New Zealand’s Land Cover Database v3.0 as published by Landcare Research in 2012). Chinn (2001) estimated the volume of ice in the Kawarau basin at 2.25 km³ using the area mean-depth relationship of the World Glacier Monitoring Service. Thus, based on Chinn’s (2001) estimate the entire water stored in the glaciers of the Kawarau could only sustain its mean flow for ~123 days, which highlights that glaciers in the catchment are less important from a hydrological perspective.”

The reviewers pointed out unexplained discrepancies in the observed and modelled hydrograph dynamics. The sentence on parameter uncertainty added is not a sufficient address to that concern.

Response C: Although we tried to account for as many uncertainty sources as possible in this study, the sources of uncertainty that were included are GCM, emission scenario, bias correction and snow model. Hence parameter uncertainty was not accounted for and can/could only be investigated as part of future work. At this point we could only speculate about the role of parameter uncertainty and therefore decided that this would go beyond the scope of this paper. The discrepancies in the hydrographs as related to inaccuracies in the modelled seasonal snow storage are already described and discussed in detail in the manuscript (P11 L11 – P12 L7).

May these differences and somewhat systematic errors be due also to the glacier melt modelling?

Response D: Very unlikely because the change in glacier volume is small between individual years when compared to the amplitude of the seasonal snow storage.

The runoff generation not well represented needs to be better assessed and clearly explained and discussed - or transparently ruled out.

Response E: There is no direct validation of runoff processes or similar due to the lack of such records/data sets. It is however very likely that the bias in the flow regime is due to the snow model as discussed in the text in quotation marks under Response C.

For the SWE-comparison of the two model variants, it would be necessary to be more transparent on how snow accumulation and melt are treated also within the glacier mass balance model component (normally, models convert snow to firn and hence 'extract' snow on glacier but not off-glacier from the water balance, snow may be differently redistributed or melt differently parameterized on or off-glacier ,etc.... - there are many potential differences that may affect the entire budget and hence the catchment-wide comparison of SWE and runoff).

Response F: As stated earlier glaciers are not very important in contributing to the generation of flow. We have added some additional information around how the glacier model works.

P4 L17 – P4 L19: “The glacier model uses three degree-day factors to calculate melt from the three storage components snow, firn and ice. After each year the remaining snow of a cell forms a new firn layer and once the firn stack reaches seven layers (as recommended by Schulla (2012)) the lowest firn layer turns into ice.”

P5 L16 – P5 L20: “Opposed to the other sub-models the glacier model was calibrated manually for the entire Clutha catchment using the annual volume estimates (1994-2010) of Willsman (2011) as calibration (1994-2001) and validation (2002-2010) source. The degree day factor (DDF) of ice (7.17 mm°C-1d-1) was based on the study of Anderson et al. (2006), while the DDF of snow (3.80 mm°C-1d-

1) was calibrated manually. The DDF of firn ($5.54 \text{ mm}^\circ\text{C}^{-1}\text{d}^{-1}$) was calculated by averaging the DDFs of snow and ice, to ensure a physically sound relationship between the three parameters (DDF of ice > DDF of firn > DDF of snow).”

If the ice melt contribution to generated runoff is too small to be relevant, the following comments may perhaps be disregarded. But if glacier ice melt is a substantial contribution (which the reader doesn't learn in the current manuscript), the necessary information on modelling must be added to the manuscript to make the study results credible and usable.

Response G: Thank you for those detailed comments around the modelling of permanent snow and ice. We have made some changes to the manuscript to avoid confusion around the estimated volume of glaciers inside the Kawarau catchment (Response B). The catchment contains 55% of the Clutha's ice cover but the overall volume is rather small in the Kawarau as well as in the other headwater sub-catchments of the Clutha. Based on the volume estimate (2.25 km³) of Chinn (2001), evenly distributed over the Kawarau catchment the ice volume corresponds to 496 mm which is small compared to the annual catchment precipitation of 2007 mm. The water stored in the glaciers of the Kawarau could only sustain its mean flow for ~123 days.

The Reviewers criticized the lack of description of the calibration procedure and the quality of the reference period simulation against which to assess future changes. Here, the dissertation by Jobst 2017 is cited, but if these details are relevant, and they are, and as this is not peer reviewed literature, the details need to be described here instead. Otherwise, it cannot be evaluated whether the quality of the model allows the conclusions drawn. Please include the calibration strategy and multi-criteria data information as it is necessary to assess how well the correct runoff generation processes are described and runoff is not generated for the wrong reason. As we all know, it is easy to model very seasonal glacial hydrographs because of the possible compensation of accumulation and melt by parameters of the hydrological model and NS efficiencies have limited value (Schaeffli and Gupta, HP - Do Nash Values have value?).

Response H: We understand that the implementation of the hydrological model is a fundamental part of accepting model results. Hence, we have tried to make the implementation and calibration process that was used to set up WaSiM clearer. Unfortunately, no data that would have allowed for the direct validation of runoff generation processes were available (e.g. groundwater table, SWE records or soil moisture) and therefore this could not be tested. We would like to stress that this is not a glacial hydrograph and it is very unlikely that glacier processes that misrepresented in the model overshadow the effect of the snow melt routine. The new figure added shows the calibration workflow that was used to calibrate the model and the calibrated parameters are explained in the figure's caption and in the added section (see below.)

P5 L5 – P5 L20: “First the parameters controlling subsurface flow were calibrated using the Nash-Sutcliffe model coefficient of efficiency (NSE). Therefore, the NSE was based on daily logarithmic streamflow values (NSElog) to account for the physical link between low flow conditions and subsurface flow components of WaSiM. The remaining parameters, controlling surface flow were calibrated using regular daily streamflow values (standard NSE) to preserve the sensitivity of the NSE to flood peaks. A regionalisation based on spatial proximity and topographical similarity was carried out next (see map in Figure 2) to parameterise sub-catchments that were ungauged or only had short records of streamflow (e.g. parameters of sub-catchment 4 → sub-catchment 3).

In the following step the two snow models were calibrated for three separate headwater sub-catchments (gauges: The Hillocks, Peat's Hut and West Wanaka as shown in Figure 1b) against

monthly streamflow (NSE_{mo}) with the Nash-Sutcliffe criterion of efficiency (NSE) as the objective function. The resulting parameter sets were then averaged resulting in a global parameter set for each of the two snow models respectively (Table 1).

Opposed to the other sub-models the glacier model was calibrated manually for the entire Clutha catchment using the annual volume estimates (1994-2010) of Willsman (2011) as calibration (1994-2001) and validation (2002-2010) source. The degree day factor (DDF) of ice ($7.17 \text{ mm}^\circ\text{C-1d-1}$) was based on the study of Anderson et al. (2006), while the DDF of snow ($3.80 \text{ mm}^\circ\text{C-1d-1}$) was calibrated manually. The DDF of firn ($5.54 \text{ mm}^\circ\text{C-1d-1}$) was calculated by averaging the $DDFs$ of snow and ice, to ensure a physically sound relationship between the three parameters (DDF of ice $>$ DDF of firn $>$ DDF of snow)."

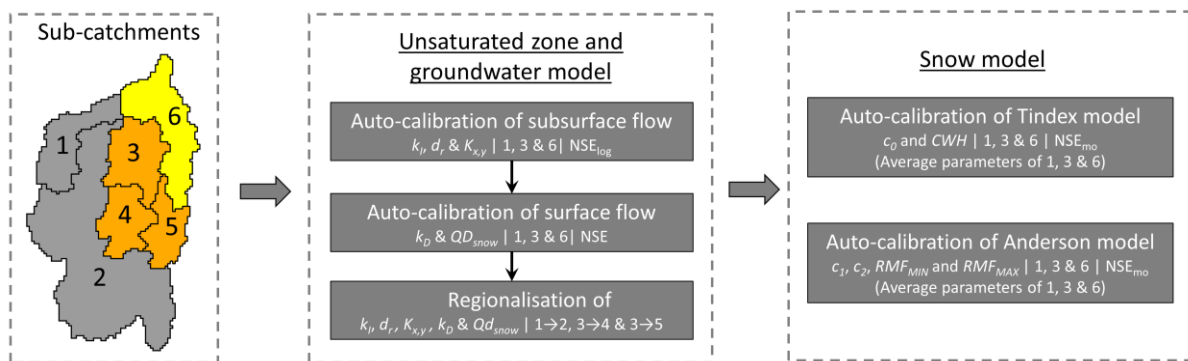


Figure 2 The calibration workflow that was used to calibrate WaSiM. For each calibration step (either manual or automatic) information is provided as follows: “calibrated parameters | calibrated sub-catchments | performance criterion used”. The parameters controlling subsurface flow encompass the drainage density for interflow (d_r), the storage coefficient of interflow (k_I) and the groundwater conductivity in x or y-direction. Surface flow is controlled by the storage coefficient of surface runoff (k_D) and the fraction of snowmelt directly becoming surface runoff (QD_{snow}). The parameters of the snow models are shown in Table 1.

Necessary information to be added also includes what happened to the glaciers between 1992 and 2012. What does that change entail for the calibrated parameters and resulting over-/underestimations in the validation period (see reviewer comments)?

Response 1: We have provided some more information on the change in the glacier volume during the reference period. The decrease from 540 mm to 265 mm during the reference period is substantial but when compared to the ice volume estimates of Willsman (2011) the model results can be considered as reasonable. Willsman (2011) found a decrease of 29% for the entire Clutha catchment between 1994 and 2010 which is less than the decrease found in this study for the Kawarau (42%). Unfortunately, there are no volume estimates for the Kawarau catchment but as shown in the figure below, which is taken from Jobst (2017) and not included in this article (number of figures is already relatively high), for the entire Clutha the modelled ice volume agrees well with the estimates of Willsman (2011) in 2010.

P6 L6 – P6 L9: “The modelled glacier storage decreased from 540 mm to 265 mm during the reference period. For comparison Willsman (2011) found a decrease of 29% for the entire Clutha catchment between 1994 and 2010, which is less than the decrease modelled here for the Kawarau (42%). As there are no other studies that have quantified the glacier volume of the Kawarau the results of this study couldn’t be assessed any further.”

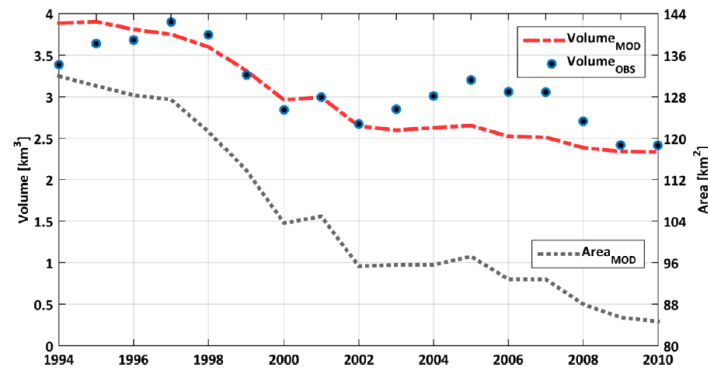


Figure 3.17 Modelled glacier volume ($\text{Volume}_{\text{MOD}}$) inside the Clutha catchment versus the volume estimates of Willsman (2011) ($\text{Volume}_{\text{OBS}}$). The modelled glacierized area (Area_{MOD}) is also displayed.

Was this change also the reason for the (to me very unusual) choice of a short recent calibration period and long past validation period? From my experience in such environments long model spin ups are required and calibration over phases of wet/cold and warm/dry phases necessary. Short calibration periods may lead to biased storage fillings and thus biased results. The motivation and procedure of these choices also need to be more transparently described as well - reader cannot be referred to read a thesis to learn this important methodological reasoning.

Response J: As stated on P4 L29-31 "The last four hydrological years of the reference period were chosen for calibration because of the higher density of weather stations compared to previous years and a better consistency of the streamflow records." The rather short calibration period was also selected to reduce the otherwise too extensive processing time of WaSiM at the relatively high spatial resolution of 1 km². Although this can be regarded as a nonstandard approach the long validation period allowed for a robust assessment of the model's performance. Hence the following explanation has been added to the manuscript.

P4 L 31 – P5 L1: "Further the relatively short calibration period constitutes a compromise between a reasonable processing time and a sufficient number of iterations (as part of the auto-calibration). Although the calibration period is rather short this means that the longer validation period allows for a robust assessment of the hydrological model's performance.

After a two-year model spin-up the individual sub-models of WaSiM..."

In addition to the question how well was the long-term glacier change in the past was modelled (see former comment that this is the best evaluation of a good snow model), it is also important to learn what happened to the glacier in the future simulations as the loss or gain will be reflected in the water balance. A future hydrograph without that information is not useful.

Response K: As mentioned above the ice volume has decreased from 540 mm to 265 mm during the reference period. It should be noted that the average annual ice volume lost is small (i.e. 14 mm) compared to the annual amplitude of the snow storage which based on WaSiM-Anderson ranges between 33 in autumn and 286 mm in spring (i.e. 253 mm).

In summary, the manuscript does not yet go far enough into modelling reasons and therefore my expectation is that it will also not have the impact it perhaps deserves. The conclusions need to go beyond the general warmer-less snow paradigm to make an interesting contribution to international science in the field. Therefore, I need to ask for further revisions regarding more transparency of the modelling details and more in-depth analyses of the reasons for the difference and range in changes founds.

Response L: The main motivation of this study was to investigate some of the key uncertainty sources in hydrological climate change impact studies for a representative headwater sub-catchment of the Clutha River (New Zealand's largest catchment). The role of the snow model was also investigated in detail and the findings are important for both alpine catchments in general and for the relatively unexplored domain of the Southern Alps. As WaSiM is a widely used hydrological model this study makes an important contribution to the existing studies that have successfully used WaSiM as a climate change impact tool.

We would like to stress the point that both reviewers had suggested minor changes but were pleased with the structure and the general content of the manuscript (Reviewer 1: "The manuscript is clearly written, and I enjoyed reading it" and Reviewer 2: "The paper is well written and comprehensive"). At this stage we believe that we have addressed the comments of the reviewers to a satisfactory degree. Adding more material around the implementation of the model would make this paper less concise (at this point the manuscript has already reached a relatively high word count of 7627). We also believe that the uncertainty analysis goes well beyond a simple "warmer, less snow and therefore more winter runoff" study. Changes in temperature and precipitation were the driving processes but we believe that the results shown here would be of interest to a wider community as this study generated some important insights around the role of the snow model and how the latter compares with the more commonly investigated uncertainty components.

Further Editorial comments

Figure 1: Please make this figure more specific to this study! Legend for the red polygon is missing. What is this? Is it relevant for this study at all? Otherwise please delete this polygon. 'snow cal' is not a nice or useful abbreviation here - please write out what is meant with that in this introductory figure. The order of a, b, c is very confusing/unusual. Please switch graphs or a, b,c to go from upper left to lower right as is usually the custom. In fact, a) is really not needed and anyway very dark colors hard to see anything. Instead, it would be more useful to give more landcover details of the modelled catchment, such as glaciers, forest vs grassland, etc. Most of what is in the map c) is not modelled and investigated in the study anyway (reservoirs and lakes only mentioned once as overall motivation).

Changed as suggested

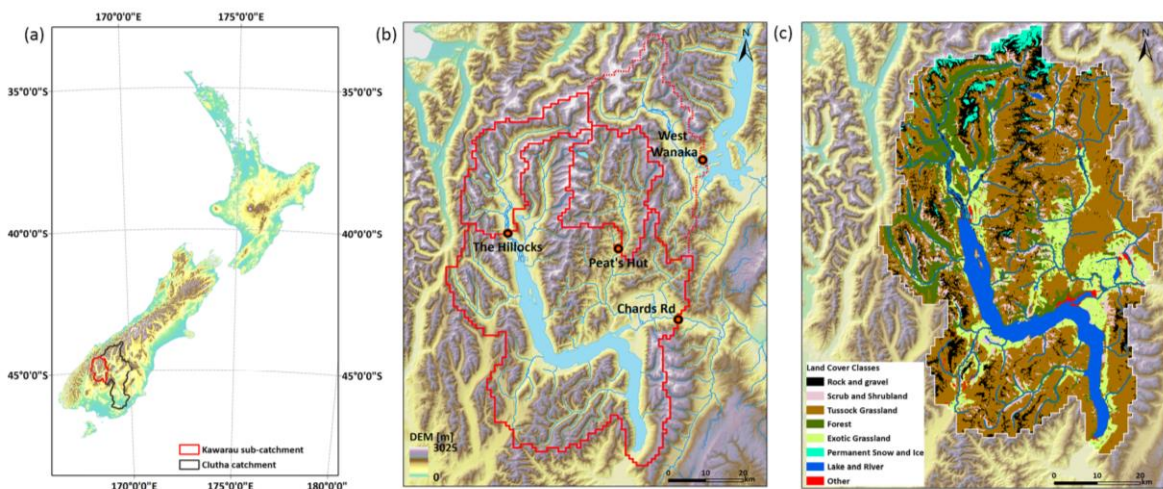


Figure 1 Maps showing (a) New Zealand with the Clutha catchment located in the lower South Island, (b) the sub-catchments of the gauges recording discharge used in this study (note that the West Wanaka sub-catchment is outside the Kawarau catchment) and (c) a land cover classification of the Kawarau sub-catchment based on New Zealand's Land Cover Database (LCDB v3.0) that was published by Landcare Research in 2012.

Figures 7-9 Axis labels are too small and too close to the axes. There is enough space to make them considerably larger and keep a bit of distance from the axes.

Changed as suggested

Figure 10a is not possible to understand without any indication what is on the x axis - what are the different ranges? The colors repeat, so it is not at all clear to me (and likely other readers) what the difference among these bars is.

Additional information has been added to the figure caption including a reference to a more detailed explanation of this figure (visualisation of uncertainty by model component) in section 3.2. This figure could potentially be omitted but considering that the radar charts are derived from this figure we would suggest not deleting it. As mentioned the figure is explained in detail in section 3.2 but essentially serves as a visual tool to portray the range of the 32 permutations for each model component.

10b-d. I suggest that these are sufficiently different to warrant a separate Figure 11 and then larger panels. At this size, it is hard to see differences clearly. Also, perhaps since the provided figure resolution is not acceptable. I assume that will be much better in the final production, but still, I think these are too small.

Changed as suggested

Table 1: give units where applicable

Changed as suggested

Table 3: why is SWE given in km³ and not also in mm similar to all other terms. This is not only highly unusual, it also doesn't allow comparison. Please change.

Changed as suggested

Intercomparison of different uncertainty sources in hydrological climate change projections for an alpine catchment (uUpper Clutha River, New Zealand)

Andreas M. Jobst¹, Daniel G. Kingston¹, Nicolas J. Cullen¹ and Josef Schmid²

5 ¹Department of Geography, University of Otago, Dunedin, PO Box 56, New Zealand

²Department of Geography, University of Munich (LMU), Munich, Germany

Correspondence to: Andreas M. Jobst (andreas.jobst@otago.ac.nz)

Abstract. As climate change is projected to alter both temperature and precipitation, snow controlled mid-latitude catchments are expected to experience substantial shifts in their seasonal regime, which will have direct implications for water management. In order to provide authoritative projections of climate change impacts, the uncertainty inherent to all components of the modelling chain needs to be accounted for. This study assesses the uncertainty in potential impacts of climate change on the hydro-climate of a headwater sub-catchment of New Zealand's largest catchment (the Clutha River) using a fully distributed hydrological model (WaSiM) and unique ensemble encompassing different uncertainty sources: General Circulation Model (GCM), emission scenario, bias correction and snow model. The inclusion of snow models is particularly important, given that (1) they are a rarely considered aspect of uncertainty in hydrological modelling studies, and (2) snow has a considerable influence on seasonal patterns of river flow in alpine catchments such as the Clutha. Projected changes in river flow for the 2050s and 2090s encompass substantial increases in streamflow from May to October, and a decline between December and March. The dominant drivers are changes in the seasonal distribution of precipitation (for the 2090s +29 to +84% in winter) and substantial decreases in the seasonal snow storage due to temperature increase. A quantitative comparison of uncertainty identified GCM structure as the dominant contributor in the seasonal streamflow signal (44-57%) followed by emission scenario (16-49%), bias correction (4-22%) and snow model (3-10%). While these findings suggest that the role of the snow model is comparatively small, its contribution to the overall uncertainty was still found to be noticeable for winter and summer.

1 Introduction

25 Over recent decades climate change has had a considerable impact on the Earth's freshwater resources (Jiménez Cisneros et al., 2014), causing, amongst others, changes in the amount of runoff (Piao et al., 2010), the timing of peak discharge (Hidalgo et al., 2009), a reduction in glacier volume (Rosenzweig et al., 2007) and an increase in flood risk (Pall et al., 2011). Future impacts under mid and late 21st century climate change are projected to intensify, affecting both the main processes and stores of the water cycle. The impacts include an increase of potential evapotranspiration (PET) over most land areas, a further

shrinkage of glaciers and changes in the runoff regime of snowmelt affected basins (Jiménez Cisneros et al., 2014). Thus 21st century climate change is expected to have substantial implications for water users and operators alike, which makes robust projections of potential changes in the seasonality and magnitude of streamflow essential.

While General Circulation Model (GCM) land surface schemes can be used for climate change impact assessments (e.g. Haddeland et al., 2011; Gudmundsson et al., 2012), their coarse resolution makes them inadequate for modelling studies at the small and meso scale. Thus, climate change impact studies typically use a cascade of models and processing steps to move between the scales of the lower resolution climate models and a separate higher resolution hydrological model (Maraun et al., 2010; Muerth et al., 2013).

As discussed by Muerth et al. (2013), the hydro-climatic model chain typically consists of the following components: emission scenario, GCM, Regional Climate Model (RCM) or statistical downscaling, bias correction and hydrological model. All of these components constitute a potential uncertainty source, and as such all need to be examined to provide a truly comprehensive understanding of the uncertainty associated with hydrological impact assessments (Teutschbein and Seibert, 2010). The uncertainty associated with the individual components of the model chain has been investigated by an increasing number of studies. Typically, GCM structure is identified as the dominant source of uncertainty (e.g. Graham et al., 2007; Prudhomme and Davies, 2009; Hagemann et al., 2011; Dobler et al., 2012). There is little agreement on the second most important source of uncertainty between the downscaling method (Wilby and Harris, 2006; Prudhomme and Davies, 2009; Dobler et al., 2012), the bias correction (Vormoor et al., 2015) or the emission scenario (Bennett et al., 2012). A common finding is that hydrological model uncertainty is less important than other uncertainty sources (i.e. GCM), but cannot be ignored (Prudhomme and Davies, 2009; Teng et al., 2012; Thompson et al., 2013; Velázquez et al., 2013). However, for certain hydrological indicators (e.g. high flow events) hydrological models can be associated with a comparable uncertainty range to the driving climate projections (e.g. Ludwig et al., 2009; Muerth et al., 2012).

As an alternative to an ensemble of different hydrological models varying in their representation of spatial variation (i.e. lumped, semi-distributed, fully distributed) and process descriptions (i.e. stochastic, conceptual or physically oriented), some studies have explored uncertainty associated with particular routines within a single model. Examples include the sensitivity of climate change impacts on the PET method used (e.g. Kay and Davies, 2008; Thompson et al., 2014). However, ~~in~~ snowmelt affected mid latitude catchments PET-related uncertainty is often relatively small (e.g. Koedyk and Kingston, 2016), with uncertainty linked to snow-related processes more important. For example, Troin et al. (2016) investigated the uncertainty introduced by the snow-melt routine in a hydrological model for three Canadian catchments. For a number of snow indicators (e.g. snow water equivalent (SWE)), most of the uncertainty was found to be caused by natural climate variability. For temporal indices (e.g. duration of snow pack) however, the different snow models showed a greater variability. Troin et al. (2016) did not look at the implications of snow model uncertainty for river flow, but the greater uncertainty associated with temporal indices could be indicative of significant implications on the timing of snowmelt and so the annual streamflow regime. Thus, the choice of the snow model as a potential uncertainty source and its implications on streamflow needs to be explored further, particularly in alpine catchments.

The aim of this present study is to investigate the contribution of the snow model and three more commonly studied uncertainty sources (i.e. GCM, emission scenario and bias correction method) to the climate change signal in hydrological projections. New Zealand's largest catchment, the Clutha, was selected for this purpose as its highly complex hydro-climate, including snow affected headwaters, makes it a particularly interesting case study. To this end, the fully distributed hydrological model WaSiM (Schulla, 2012) was implemented for an alpine sub-catchment of the Clutha, with a total of 32 separate hydrological simulations produced. These comprised two emissions scenarios, four GCMs, two bias correction methods and two snow models. Although previous New Zealand studies (including for the Clutha) have examined multiple GCM scenarios (e.g. Poyck et al., 2011; Gawith et al., 2012; Caruso et al., 2016), none have used an ensemble covering the present range of uncertainty sources. Furthermore, in using WaSiM this will be the first application of a fully distributed and grid-based hydrological model for this purpose in a medium to large-scale New Zealand catchment. Consequently, this study will generate the most complete assessment of climate change impacts on ~~stream~~river-flow and associated uncertainty for an alpine New Zealand catchment. Importantly, the study will also speak more widely to the issue of snowmelt uncertainty under climate change in alpine catchments.

2 Data and methods

15 2.1 The study area

The Clutha/Mata-Au is the largest catchment (20586 km²) in New Zealand and is situated in the lower half of the South Island, extending eastwards from the Southern Alps (Figure 1a). It has the highest average streamflow of any river in New Zealand (approximately 570 m³ s⁻¹) and drains 6% of the South Island's water (Murray, 1975). The catchment is characterized by a highly variable hydro-climate ranging from very humid alpine headwaters dominated by seasonal snow accumulation and melt, to substantially drier areas in the central catchment. The Clutha catchment can thus be considered broadly representative of most of the South Island's hydrologic and climatic domain, and so an ideal candidate for investigating climate change impacts.

As described in Jobst (2017) WaSiM was implemented for the entire Clutha catchment as a tool for climate change impact modelling. Most of the upper and lower Clutha catchment are under extensive water management (particularly for hydro-electric dams and water abstractions), except for the north-western part (gauge Chards Rd in Figure 1b), which is characterised by natural flow conditions. As the focus of this climate change impact study is on potential changes in natural streamflow and seasonal snow, the results focus mostly on the Kawarau River sub-catchment (~~Chards Rd gauge~~) only (see Jobst (2017) for model simulations at the other key sites of the Clutha). The catchment area of the Kawarau River (at Chards Rd) is 4541 km² (22% of entire Clutha basin) and with a mean discharge of 212 m³ s⁻¹ is the largest component of the upper Clutha basin, comprising 36% of flow at the catchment outlet. Snowmelt contributes approximately 20% of annual flow. Streamflow at Chards Rd is also highly correlated with the headwater sub-catchment of Lake Wanaka (located to the east of the Kawarau sub-catchment), while the outflow of Lake Hawea is controlled by a dam for hydropower generation further downstream.

Most of the Kawarau sub-catchment is covered by indigenous tussock grassland followed by low producing exotic grassland and ~~indigenous-forest~~ (Figure 1c). The elevation of the sub-catchment ranges between 300 and 2800 m, ~~with an ice cover of approximately 84 km² which corresponds to 55% of the Clutha catchment's ice cover (New Zealand's Land Cover Database v3.0 as published by Landcare Research in 2012).~~ The ice-covered area inside the Kawarau catchment is small and based on 2001/2002 satellite imagery amounts to 1.8% or 84 km² (New Zealand's Land Cover Database v3.0 as published by Landcare Research in 2012). Chinn (2001) estimated the volume of ice in the Kawarau basin at 2.25 km³ using the area mean-depth relationship of the World Glacier Monitoring Service. Thus, based on Chinn's (2001) estimate the water stored in the glaciers of the Kawarau could only sustain its mean flow for ~123 days, which highlights that glaciers in the catchment are less important from a hydrological perspective.

10 2.2 The WaSiM model of the Clutha

The fully distributed and physically-oriented hydrological model WaSiM-Richards (version 9.06.10) was implemented at a spatial resolution of 1 km and at a daily time step. The main components of this implementation of WaSiM are described briefly here – for a more detailed description see Schulla (2012). The modelling of PET is solved by the Penman Monteith approach, while actual evapotranspiration (ET) is a function of the simulated soil water content. Soil and groundwater processes are described by finite differencing of the 1D-Richards equation combined with a 2D groundwater model. In addition, WaSiM's dynamic glacier model was used to describe the glacial processes for the ice-covered cells located in the upper catchment. The glacier model uses three degree-day factors to calculate melt from the three storage components snow, firn and ice. After each year the remaining snow of a cell forms a new firn layer and once the firn stack reaches seven layers (as recommended by Schulla (2012)) the lowest firn layer turns into ice.

20 WaSiM was parameterised using both remotely sensed data (i.e. MODIS-15A2-1km for leaf area index) and values obtained from the literature. Two versions of this WaSiM implementation were set up, one with a simple temperature index (Tindex) snow-melt routine (Schulla, 2012) and the other with the conceptual energy balance model of Anderson (1973). The Tindex model calculates the melting rate via a degree-day factor multiplied by the difference between actual temperature and the melting point temperature. The Anderson model is more complex as it computes four separate melt fractions, ~~and~~ accounts for radiation by using a seasonal melt factor and. ~~Further the Anderson approach~~ also models the refreezing of liquid water stored in the snow pack when the actual temperature is below melting point.

Station-based meteorological observations of mean daily air temperature (Tmean), precipitation, solar radiation, relative humidity and wind speed were interpolated (Jobst, 2017; Jobst et al., 2017) and served as input to WaSiM during the calibration (2008-2012) and validation (1992-2008) periods. The last four hydrological years of the reference period were chosen for calibration because of the higher density of weather stations compared to previous years and a better consistency of the streamflow records. Further the relatively short calibration period constitutes a compromise between a reasonable processing time and a sufficient number of iterations (as part of the auto-calibration). Although the calibration period is rather short this means that the longer validation period allows for a robust assessment of the hydrological model's performance.

After a two-year model spin-up the individual sub-models of WaSiM (unsaturated zone, groundwater, snow and glacier model) were calibrated iteratively using a combination of auto-calibration and manual parameter optimization (Figure 2; see Jobst (2017) for a more detailed description of the calibration process). Particle swarm optimization (PSO) (Kennedy and Eberhart, 1995) was used for auto-calibration due to its effective performance during the first iterations and fast operation (Jiang et al., 2010), allowing for an adequate compromise between processing time and efficiency. First the parameters controlling subsurface flow were calibrated using the Nash-Sutcliffe model coefficient of efficiency (NSE). Therefore, the NSE was based on daily logarithmic streamflow values (NSE_{log}) to account for the physical link between low flow conditions and subsurface flow components of WaSiM. The remaining parameters, controlling surface flow were calibrated using regular daily streamflow values (standard NSE) to preserve the sensitivity of the NSE to flood peaks. A regionalisation based on spatial proximity and topographical similarity was carried out next (see map in Figure 2) to parameterise sub-catchments that were ungauged or only had short records of streamflow (e.g. parameters of sub-catchment 4 → sub-catchment 3).

In the following step the two snow models were calibrated for three separate headwater sub-catchments (gauges: The Hillocks, Peat's Hut and West Wanaka as shown in Figure 1be) against monthly streamflow (NSE_{mo}) with the Nash-Sutcliffe criterion of efficiency (NSE) as the objective function. The resulting parameter sets were then averaged resulting in a global parameter set for each of the two snow models respectively (Table 1).

Opposed to the other sub-models the glacier model was calibrated manually for the entire Clutha catchment using the annual volume estimates (1994-2010) of Willsman (2011) as calibration (1994-2001) and validation (2002-2010) source. The degree day factor (DDF) of ice ($7.17 \text{ mm}^\circ\text{C}^{-1}\text{d}^{-1}$) was based on the study of Anderson et al. (2006), while the DDF of snow ($3.80 \text{ mm}^\circ\text{C}^{-1}\text{d}^{-1}$) was calibrated manually. The DDF of firn ($5.54 \text{ mm}^\circ\text{C}^{-1}\text{d}^{-1}$) was calculated by averaging the DDFs of snow and ice, to ensure a physically sound relationship between the three parameters (DDF of ice > DDF of firn > DDF of snow).

With regards to the validation of streamflow while daily NSE values were lower indicated a reasonable performance for the three tributaries (Dart, Shotover and Matukituki River), while monthly NSE values indicated a good performance at these sites (Table 2). For the Matukituki River the validation of both WaSiM versions revealed a substantially better simulation of monthly streamflow (NSE of 0.83 and 0.82, respectively) when compared to the TopNet based modelling study of Gawith et al. (2012) (NSE of 0.68). For Chards Rd the validation of WaSiM-Anderson and WaSiM-Tindex revealed a strong performance at the daily and monthly time scale, with NSE values between 0.85 and 0.90 across all model versions, timescales and time periods (Table 2).— The monthly hydrographs of WaSiM-Anderson and WaSiM-Tindex (Figure 43c) further indicate a realistic representation of observed daily runoff at Chards Rd (only the first four years of the validation period are shown for clarity). Obvious inaccuracies of both WaSiM versions are an underestimation of larger flow events during the melt period (e.g. November-December-January 1995, 1996 and 2000) and an overestimation during autumn (e.g. April-May 1994 and 1996). The likeliest explanation is that not enough snow is being accumulated from autumn to early winter and consequently the main melt peaks are under-simulated. This is largely supported by the monthly SWE simulations (Figure 4b) and can be exemplified for the snow accumulation period of 1992 where WaSiM-Anderson results in a greater accumulation of SWE compared to WaSiM-Tindex. In the following melt period (November 1992 to February 1993) the observed peak is then approximated

better by WaSiM-Anderson due to the larger SWE storage. Overall the inaccuracies are generally larger with WaSiM-Tindex as can be seen in December 1994 (Figure 43). Compared to WaSiM-Anderson the snow pack melts slower which causes a greater underestimation of runoff during the main melt period. Overall the visually better good performance of WaSiM-Anderson for the 1992-1996 period is substantiated by the monthly daily NSE values which correspond to 0.940.89 and 0.870.88 for WaSiM-

5 Anderson and WaSiM-Tindex, respectively.

The modelled glacier storage decreased from 540 mm to 265 mm during the reference period. For comparison Willman (2011) found a decrease of 29% for the entire Clutha catchment between 1994 and 2010, which is less than the decrease modelled here for the Kawarau (42%). As there are no other studies that have quantified the glacier volume of the Kawarau the results of this study couldn't be assessed any further.

10 Due to the high annual precipitation totals the grid cells in the upper catchment could potentially be prone to large build-ups of snow that persist throughout the melting period and into the next hydrological year. As discussed by Freudiger et al. (2017) such snow towers can cause substantial errors in the modelled water balance and should therefore be made transparent. As indicated in Figure 4b the snow pack of some cells persists across all of the melt seasons with SWE values at the end of the melt season ranging between 7 and 46 mm (21 and 77 mm) for WaSiM- Anderson (WaSiM-Tindex). To further investigate the significance of this potential issue, all cells where the SWE value at the end of the hydrological year exceeded the annual snowfall (excluding glacier cells), were marked. The combined relative area of these cells is shown in Figure 4a and does not exceed 1% for WaSiM-Anderson, while for WaSiM-Tindex the affected area does not exceed 2% except for the years 1997 and 1998. The mean exceedance (SWE/snowfall) for the individual years based on all affected cells varies from 1.3 to 1.8 for WaSiM-Anderson and from 1.4 to 2.3 for WaSiM-Tindex. Thus, for a number of cells along the main divide the modelled

15 build-ups of snow could be classified as potential snow towers. However, the combined area of these cells is small, which means that the effect on the modelled water balance is negligible.

20

2.3 The model cascade

Most existing impact studies in the New Zealand domain (Poyck et al., 2011; Srinivasan et al., 2011; Zammit and Woods, 2011; Zemansky et al., 2012) have been based on statistically downscaled GCM simulations provided by the National Institute

25 of Water and Atmospheric Research (NIWA) (Ministry for the Environment, 2008). More recently a small ensemble of four GCMs (CM2.1-GFDL, ECHAM5, HadCM3 and MK3.5-CSIRO) based on the A1B and A2 SRES emissions scenarios has been dynamically downscaled for the New Zealand domain using the HadRM3P RCM (Ackerley et al., 2012), and it is this ensemble of eight dynamically downscaled GCM simulations that forms the data set for the current study.

A model chain was constructed (Figure 32) to process the raw RCM runs (from 1990 to 2099) and generate high resolution

30 climate change projections at the hydrological model scale. Two different bias correction methods, linear transformation (LT; as described in Lenderink et al. (2007)) and quantile mapping (QM; as described in Mpelasoka and Chiew (2009)), were used to correct the RCM data. Both methods have been successfully used by a number of studies (e.g. Boé et al., 2007; Chen et al.,

2013; Gutjahr and Heinemann, 2013) and were selected here to allow for a direct comparison between a simple correction method based on additive or multiplicative correction terms (LT) and the more complex distribution-based QM approach.

To bridge the gap between the RCM grid (~27 km) and the hydrological model grid (1 km) an additional statistical downscaling step was required. The downscaling of precipitation (and the remaining three variables) is based on the topographical scaling approach of Frueh et al. (2006), while maximum and minimum air temperature are scaled via monthly lapse rate models (as described in Jobst et al. (2017) but excluding the thin plate spline layer). As part of the downscaling, additional processing steps were adopted from Marke (2011) to ensure the conservation of mass and energy when transforming the RCM data between the model scales.

3. Results

10 3.1 Baseline simulations

For the historic analysis, the ensemble was divided into four sub-ensembles composed of the two bias correction methods and the two snow models (i.e. QM-Anderson, QM-Tindex, LT-Anderson and LT-Tindex). The regimes (i.e. monthly runoff values averaged across all years) of the eight RCM driven simulations of each sub-ensemble were compared to the observed runoff regime (OBS) and the modelled runoff regime forced by the observed meteorology (MOD-METEO_{OBS}).

15 The skill in reproducing the observed historic regime varies substantially depending on both the bias correction method and the snowmelt routine (Figure 54). Overall QM-Anderson gave the most realistic approximation of the observed regime, although still with some overestimation in May (late autumn) followed by an underestimation during July and August (winter). QM-Tindex and LT-Anderson also underestimate the main peak, however the general fit of their RCM members is still relatively close to the observed regime. The largest discrepancies occurred with LT-Tindex, with a substantially flatter regime, 20 mainly due to too much flow being generated between May and September leading to an underestimation of the main peak (November to January). Overall the LT method shows a lower skill in reproducing the observed regime, which is especially pronounced in combination with Tindex. This behaviour points to a high sensitivity of the modelled regime towards the bias correction method and generally speaking the meteorological forcing.

The RCM driven runs agree more closely with MOD-METEO_{OBS} than with the observed regime, as monthly over- and 25 underestimations of MOD-METEO_{OBS} have propagated into the RCM driven WaSiM simulations. This was expected as the RCM climate data have been tuned (i.e. bias corrected) to the station-interpolated meteorology that was used to drive MOD-METEO_{OBS}.

Regarding the water balance (Table 3), the observed annual precipitation of the Kawarau sub-catchment (2007 mm) was underestimated by both the QM (1926 mm) and the LT sub-ensemble (1931 mm) during the reference period. A small part of 30 that difference is caused by the shorter 360-day calendar of the RCM runs. Compared to MOD-METEO_{OBS}, ET was modelled almost identically by QM-Anderson and QM-Tindex, with slightly larger discrepancies (-1%) under LT-Anderson and LT-

Tindex. Regarding runoff, QM and LT resulted in an underestimation of -6 and -5%, respectively, while the choice of the snow model had only a negligible impact.

In terms of the seasonal SWE volume, the QM-Anderson runs agree more closely with MOD-METEO_{OBS} than the other sub-ensembles. The differences in the seasonal SWE volumes range from 3% in spring to 6% in autumn. Compared to MOD-METEO_{OBS} the modelled SWE volumes of LT-Tindex were almost identical for summer and autumn but substantially lower for winter (-18%) and spring (-16%). Thus, the poor agreement between the observed runoff regime and the LT-Tindex runs (Figure 45) can very likely be explained by too much melt being modelled between winter and early spring. The latter results in a reduced SWE volume, which is insufficient to supply streamflow with enough snowmelt during late spring and summer.

3.2 The climate change signals of precipitation and air temperature

10 The climate change signals of Tmean and precipitation that are presented in the following section correspond to the mean change of the spatio-temporal average between a future period (either 2050s or 2090s) and the reference period. For precipitation, the spread of the 2050s summer climate change signal (Figure 6-5a, left box) is almost completely caused by the GCM structure. Both the emission scenario and bias correction method have negligible effects on the extent of the signal range and median, with the latter showing a near zero change in precipitation. Regarding the 2090s summer, the median change is more negative, while both the emission scenario and bias correction cause a slight increase in the uncertainty range. A different situation can be seen for the 2050s winter (Figure 56e), where the extent of the range is largely determined by the emission scenario. For the 2090s winter, all three components have a considerable impact on the uncertainty range. Here, the GCM spread is the largest of all seasons and future periods. It can also be seen that the precipitation signal is noticeably higher for the A2 sub-ensemble (mainly caused by ECHAM5-A2). In addition, the selection of the bias correction method considerably increases the extent of the whole ensemble, resulting in a total uncertainty range spanning 55.3 percentage points (i.e. a 28.5% to 83.8% increase from the baseline).

For all seasons the uncertainty in the Tmean signals during the 2050s is predominantly caused by the GCM structure (Figure 67). The selection of the emission scenario becomes a major source of uncertainty in the 2090s with most of the A2 members projecting a stronger signal than their corresponding A1B members. However, this only holds for members stemming from the same GCM (e.g. ECHAM5-A1B and ECHAM5-A2), as can be seen for the 2090s winter, where an A1B member (MK3.5-CSIRO) has a greater warming signal than two of the A2 members (HadCM3 and CM2.1-GFDL).

3.3 The hydrological signals

3.3.1 Runoff

For both future periods the historic melt-driven December peak in the annual regime is projected to move earlier in the year (Figure 78). In the 2050s, the highest monthly mean flow is projected to occur between October and November, with a further

shift for the 2090s (to September and October). The most striking transformation for the 2090s is the dramatic enhancement of monthly flows during winter and spring, with uninterrupted increases from May to October.

In order to specifically compare the contribution of the snow model with the remaining sources of uncertainty, the seasonal signals in runoff are shown separately for WaSiM-Anderson and WaSiM-Tindex (Figure 89). It can be seen that the influence of the snowmelt routine on seasonal flows is comparatively small for both periods and during all seasons. The most noticeable difference is an enhancement of the decrease during summer (Figure 8a, e) and a more pronounced increase during winter (Figure 8c, g) when using the Anderson model. Compared to the snow model the effect of the bias correction on the overall spread is more important. Positive signals were found to be enhanced by the QM method and vice versa for negative signals. Further, the influence of bias correction becomes visibly more important in the 2090s (except for autumn).

While both the bias correction and the snow model contribute to the overall spread, the GCM and the choice of the emission scenario appear to be the dominant sources of uncertainty. In most seasons the GCM range differs substantially depending on the underlying emission scenario. In the 2050s period this becomes especially apparent for the autumn season (Figure 89ab), during which the A2 runs show a much greater spread than the A1B runs. Regarding the 2090s period (Figure 89be-h), differences between the A2 and the A1B runs become more pronounced and the most extreme signals are all represented by an A2 member.

Regarding the climate change signal derived from the overall ensemble mean, the largest changes were projected for winter and summer. For the 2090s (2050s) summer runoff is projected to decrease by -24% (-10%), with a substantial increase of 71% (29%) during winter. The overall ensemble spread becomes largest for the 2090s winter season, ranging from 40 to 116%.

3.3.2 Snow water equivalent

~~The future simulations of the monthly SWE storage averaged over the Kawarau sub-catchment (gauge Chards Rd) are depicted in Figure 9 for the four sub-ensembles. During the historic period~~ The historic simulations of the monthly SWE storage volumes (blue envelope) varied considerably for the four sub-ensembles (as expected from the seasonal values in Table 3), with the lowest and highest volumes modelled by LT-Tindex and QM-Anderson, respectively. Despite these differences in the historic simulations envelope, the relative changes are similar between the four sub-ensembles and thus the results are primarily discussed only shown for the QM-Anderson sub-ensemble (Figure 10). A general observation is that the larger spread in the precipitation signal of the QM runs (Figure 56) has clearly propagated into the uncertainty range of the monthly SWE volume, while the LT envelopes are visibly narrower during both time periods (particularly for the 2090s).

For the 2050s period (Figure 9a10a), the A1B envelope predominantly lies within the upper and lower bounds of the A2 envelope, with the latter showing a substantially greater spread. The proportion of the two envelopes becomes reversed in the 2090s period, when the A1B envelope surpasses the A2 envelope in all months (Figure 109be).

Although the 2090s envelopes of the four sub-ensembles have a relatively large overlapping area during winter and spring, all of the A2 members have a lower SWE volume than their A1B counterparts. It is however noticeable that the A1B member MK3.5-CSIRO has a lower SWE volume than the ECHAM5 A2 member, which is associated with a greater warming signal

during winter and spring (Figure 67). A closer inspection of the corresponding climate signals revealed that the MK3.5-CSIRO-A1B signal can be primarily attributed to less precipitation (second smallest increase in winter precipitation of the ensemble), which combined with a relatively strong warming signal (strongest of all A1B members) would have resulted in less snowfall and therefore less snow accumulating. This indicates that despite the warming signal being a key driver of SWE changes in the future, the precipitation signal also plays an important role, which in this case led to an enhancement of the negative SWE signal.

As shown (Figure 11) by the transient simulations of mean SWE from July to December (months with historically the highest SWE) there is no clear distinction between the A1B and A2 simulations until the last quarter of the 21st century when the median of the A2 runs stays consistently under the A1B runs. While the median SWE based on all 16 QM runs decreases substantially (60%) during the 110-year period, years with large accumulations of SWE still occur in the second half of the 21st century for individual members and years. The negative trend was found to be more pronounced for the glaciers in the catchment, which are projected to lose 93% of their volume (91% based on A1B and 94% based on A2) by 2099.

3.4 Quantifying the uncertainty in the seasonal runoff signal

In order to quantify the uncertainty induced by the individual components of the model chain compared to the overall uncertainty in runoff projections, the approach of Muerth et al. (2012) was adopted. First, the approach is exemplified for the uncertainty quantification of the winter runoff signal (Figure 10a12). An uncertainty component (e.g. GCM) is selected and all possible permutations between the selected and the remaining three model components are computed, resulting in 32 combinations ($4 \cdot 2 \cdot 2 \cdot 2$). In the next step, the currently selected component is varied (four GCMs = four circles), while the other three components (emission scenario, bias correction and hydrological model) are fixed to a certain combination of their members. As such, all of the four circles spanning the first bar (Figure 10a12, left, first segment) have the emission scenario fixed to A1B (orange quarter), the bias correction method fixed to QM (white quarter) and the snow model fixed to Anderson (blue quarter), while the fourth quarter, which corresponds to one of the four GCM members (CM2.1-GFDL, ECHAM5, HadCM3 and MK3.5-CSIRO), is varied. Thus, the effect of the GCM has been isolated by fixing the other components to one particular combination. This step is repeated for all the possible combinations between the three remaining components, translating to a total of eight combinations (eight bars). As each of the eight bars in the GCM segment (Figure 10a12, first segment) contains four circles, all 32 possible permutations have been accounted for. The same procedure is then repeated for the remaining three uncertainty components. The mean contributions to the overall uncertainty (including the standard deviations) of the individual components are then displayed in a radar chart (Figure 130ab).

The uncertainty analysis (Figure 130ab-de) identified the GCM as the primary source of uncertainty across all seasons (44-57% change in runoff). The selection of the emission scenario was the second largest contributor (16-49%), except for winter when the choice of bias correction was greater (22% vs. 16%). The uncertainty induced by the emission scenario showed a pronounced seasonal variation and was found to be largest during summer (33%) and autumn (49%). A likely explanation for the latter are the significantly different Tmean signals under A1B and A2 (Figure 67), which translate to different ET rates and

consequently variable changes in runoff. This is supported by the fact that the most extreme decreases in summer and autumn runoff occurred under the A2 scenario (not shown here). The contribution of bias correction to the overall uncertainty ranged from 4% in autumn to 22% in winter and was higher (except for autumn) than the relative contribution of the snow model (3-10%).

- 5 As described in Muerth et al. (2012) the standard deviation associated with the relative uncertainty contribution of an individual component indicates its degree of dependence on the other model components. Here the standard deviation was clearly largest for the emission scenario and the GCM. The standard deviations of both components also varied seasonally and were found to be largest during spring and autumn (Figure 130be, de). Thus, it can be stated that, during spring and autumn the uncertainty induced by the GCM (same holds for emission scenario) is associated with a relatively large dependence on the other variables.

10 4. Discussion

Before the climate change uncertainty assessment was carried out, the hydrological simulations were analysed during the reference period. Performance in reproducing the observed regime varied depending on both the selection of the snow model and the bias correction method. The bias correction method was expected to only have a minor effect on the simulated monthly runoff during the reference period, which made the observed sensitivity somewhat unexpected. A potential explanation could be related to the predominant air temperature during mature precipitation events along the main divide. At Brewster Glacier (located just outside the Clutha catchment, along the main divide west of Lake Hawea – Figure 1), Cullen and Conway (2015) found air temperature to be frequently around the rain-snow threshold during events with major solid precipitation, which led to the conclusion that the accumulation of snow in areas along the main divide is vulnerable to small variations in air temperature. The relatively large variability during the reference period was therefore seen as a first indicator for a potentially high sensitivity of the modelled snow storage and runoff to projected warming. The historic analysis also showed that the observed regime was captured more realistically by WaSiM-Anderson opposed to WaSiM-Tindex. Studies targeting the controlling processes of snowmelt in the Southern Alps (Prowse and Owens, 1982; Sims and Orwin, 2011) identified net radiation as an important driver of snowmelt (in addition to sensible heat). Thus, the better performance of the conceptual energy balance method (WaSiM-Anderson) compared to the Tindex model can likely be explained by the advances of accounting for individual melt fractions and using a seasonal radiation melt factor (see Anderson (1973) or Schulla (2012)). Overall the baseline analysis showed that the individual sub-ensembles performed differently and that the observed regime was not always enveloped by the corresponding range of simulations. This introduces some additional uncertainty into the projections that could not be quantified or accounted for in this study. A potential explanation is that neither snow model was able to accurately represent all of the spatio-temporal variation in the snowmelt process across the catchment, and that some driving processes (i.e. radiation induced events) are also not represented adequately in either snow model. Either improved empirical relationships or a greater physical component to snowmelt modelling would be beneficial in this respect for future research. [Accounting for snow redistribution by avalanches or wind as described in Freudiger et al. \(2017\) should also be](#)

investigated as this would potentially reduce some of the relatively large accumulations of snow that were found to persist over multiple years on some cells in the upper catchment during the reference period.-Inaccuracies in the meteorological fields

(METEO_{OBS}) that were used for the bias correction could also have caused some of the seasonal over- and underestimations in the hydrological regime. As discussed in Jobst et al. (2017) the climate network in the upper Clutha is sparse with very few sites located in medium to high elevations. Notwithstanding the improved representation of temperature provided by the Jobst et al. (2017) dataset compared to other products, the remaining biases in this temperature field would have also propagated into the bias corrected RCM fields and the corresponding hydrological baseline simulations.

For the two future periods (2050s and 2090s) the projections revealed substantial increases in runoff from May to October, and a decline between November and March. The dominant drivers behind this regime shift were changes in the seasonal distribution of precipitation (for the 2090s-winter +29 to +84%) and a rise in air temperature causing decreases in the seasonal snow storage. These findings are mostly consistent with previous New Zealand based climate change impact assessments. In a New Zealand wide study Hendriks et al. (2012) also modelled substantial reductions in the peak snow accumulations along the Southern Alps, which they attributed to decreases in the fraction of solid precipitation due to increases in air temperature. Using a semi-distributed hydrological model (TopNet) and ensemble of 12 CMIP-3 GCMs (including the four used herein), Poyck et al. (2011) and Gawith et al. (2012) found a similar ensemble-mean increase in winter streamflow in the 2090s (for Balclutha and the upper Clutha – Matukituki catchment), although only relatively small decreases in summer river flow. In the upper Waitaki catchment (9490 km², located north-east of the Clutha and also with headwaters bordering the Main Divide of the Southern Alps), Caruso et al. (2017) found comparable large increases in lake inflows during winter (i.e. 76% for August) and a noticeable decrease in summer (i.e. -13% for February) using the same hydrological model and GCM ensemble as Poyck et al. (2011). An increase in winter precipitation was also identified as the main driver for the Waitaki.

Globally, similar changes in streamflow have been reported for many alpine catchments, for example in British Columbia (Mandal and Simonovic, 2017), Oregon (Chang and Jung, 2010) and the Austrian Alps (Laghari et al., 2012). In addition to increased winter precipitation, a reduction in solid precipitation is often reported to lead to an earlier melt peak and further enhanced winter flow (Kundzewicz, 2008). Here, a decrease in the proportion of solid precipitation combined with an intensification of snowmelt was also found to contribute to more flow being generated during winter and spring, but the main driver remained the increase in winter precipitation.

Analysis of the uncertainty in the hydrological projections for the upper Clutha (Figure 12 and 13) showed that although the total spread of hydrological projections was large (i.e. increase of 40-116% for winter), for most seasons (except autumn) the direction of change was found to be consistent amongst individual members (increases in winter and spring, decreases in summer). The main contributors to the spread in the projections for seasonal flow were (in ascending order): snow model (3-10%), bias correction method (4-22%), emission scenario (16-49%) and GCM (44-57%). As in this study, a large body of existing hydrological impact studies also identified GCM structure as the dominant source of uncertainty (e.g. Kingston and Taylor, 2010; Hughes et al., 2011; Teng et al., 2012; Thompson et al., 2014). It should be noted that the four GCMs constitute a subset of a total of 12 GCMs which had been previously selected by NIWA on the basis of a performance assessment for the

South-Pacific region (Ministry for the Environment, 2008). In terms of Tmean signal (A1B) the four GCMs had the 5th, 10th, 11th and 12th highest warming– hence the A1B projections used in this study are at the lower end of the “full” GCM envelope.

A large part of the GCM related uncertainty was found to be caused by the precipitation signal, which became especially uncertain during the winter season. This finding is in agreement with a number of studies targeting alpine catchments such as the Hindu-Kush-Himalayan region (Palazzi et al., 2014; Lutz et al., 2016), the Pacific Northwest of the US (Jung et al., 2012), Western Oregon (Chang and Jung, 2010) and the Southern Alps of New Zealand (Zammit and Woods, 2011). Hence constraining and accounting for the uncertainty associated with the precipitation output of GCMs and RCMs remains a major research challenge in hydrological impact studies.

Emission uncertainty was identified as the second most important source during most seasons, while in winter bias correction was found to introduce a similar level of uncertainty. These findings generally agree with the study of Prudhomme and Davies (2009), in which emission scenario and downscaling (RCM vs. statistical method) uncertainty were of a comparable magnitude, but still considerably smaller when compared to GCM uncertainty. For alpine catchments in British Columbia the ranking order of uncertainty sources computed by Bennett et al. (2012) was also led by the GCM, followed by the emission scenario and in third hydrological parameter uncertainty.

As described in Kay and Davies (2008) and Thompson et al. (2014), different versions of the same hydrological model can be developed that differ in one particular routine (i.e. PET) allowing for a process specific uncertainty analysis. In the upper Clutha catchment, the high precipitation intensity in the headwaters combined with the relatively high proportion of snowmelt (~20%) means that the seasonal regime is largely controlled by the process of snowmelt rather than PET, which made the upper Clutha an appropriate candidate for the snow model specific uncertainty analysis. By using the two WaSiM versions that only differ in their snowmelt routine the contribution of that process to the overall uncertainty could be assessed in isolation.

As expected, the contribution of the snow model was highest for winter (10%). However interestingly, the contribution of the snow model was still relatively high during summer (8%), a time of year when the influence of melt processes on streamflow were expected to be minor. This can likely be explained by the larger SWE volume that was modelled by Tindex (compared to Anderson) during the summer reference period (Table 3). Thus, the Tindex SWE storage had the potential to release more melt water (compared to baseline) under the projected warming, which translated into an attenuation of the decrease in summer runoff. This is supported by the fact that the negative changes in summer runoff are consistently less pronounced for the Tindex model (not shown here). For autumn and spring, the snow model only added a small proportion (5 and 3%, respectively) to the overall uncertainty. Considering that spring is (historically) the main melt period, projections were expected to vary more depending on the choice of the snow model. Hence for the spring season the results suggest that under the projected warming the effect of the snow model can be considered negligible especially when compared to the GCM and the emission scenario. At 10% of overall uncertainty in winter, the effect of the snow model is noticeable but substantially smaller than the variation caused by the GCM output (48%) (uncertainty of bias correction and emission scenario corresponds to 22 and 16%, respectively).

The study of Troin et al. (2016), which focused on the direct output of the snow model (i.e. SWE or duration of snow pack), came to comparable conclusions in the sense that hydrological models are not the major source of uncertainty for SWE projections. In their study, natural variability had a far greater effect on the projections for the individual snow indicators as the snow model component, which was shown here in a similar way for GCM structure.

5 5. Conclusions

The implementation of WaSiM for the Clutha River constitutes the first application of a fully distributed and grid-based hydrological model for climate change impact assessment in a large scale New Zealand catchment. The model chain that was built here to force WaSiM with RCM simulations can be regarded as an important contribution to the existing body of climate change impact studies targeting snowmelt affected mid-latitude alpine catchments. The projections for the end of the 21st century encompass substantial increases in runoff for winter (71%) and spring (35%), while summer runoff is projected to decrease (-24%). The key drivers behind the changes in the regime were found to be an increase in winter precipitation and a reduction in the SWE storage between winter and spring. The changes in the regime will likely impact the capacity of the two hydropower schemes (Clyde and Roxburgh; Figure 1) located downstream of Chards Rd, where production can be expected to increase from winter to early ~~spring, and~~ spring and decrease during the summer months.

Adopting the approach of Muerth et al. (2012), this study allowed the contribution of the individual uncertainty sources to be quantified in a more objective way opposed to a mere visual interpretation of results. For the first time the role of the rarely investigated snowmelt routine was explored together with three of the key uncertainty sources in hydrological impact studies (i.e. GCM, emission scenario and bias correction method). While all components contributed to the total ensemble uncertainty, the selection of the GCM introduced the biggest spread to the range of runoff projections during all seasons. When looking at the climate signals (Figure ~~65~~ and ~~76~~) it becomes obvious that the uncertainty stemming from the precipitation signal (especially during winter) is the primary driver behind the large uncertainty in the hydrological projections and should therefore be the focus of future studies. In this context combining the limited number of RCM simulations with sophisticated statistical techniques (e.g. the use of probability density functions as described in Tait et al. (2016)) could help to more fully explore the uncertainty range associated with the precipitation signal. Further, since the completion of this study additional RCM simulations based on RCP (Representative Concentration Pathways) scenarios and CMIP-5 GCMs have been generated for the New Zealand domain (Ministry for the Environment, 2016), which could be used to enlarge the existing ensemble of hydro-climatic projections for the Clutha.

The uncertainty linked to the snow model, which showed a pronounced seasonal variation (ranging from 3% in spring to 10% in winter), was found to be smaller when compared to the other components, but the findings of this study suggest that it should not be ignored as its effect was shown to be significant for both winter and summer runoff. Another important finding from this study is that the contribution of the snow model and the other model components to the overall uncertainty possesses a high inter-annual variability. While there was consistency regarding the main uncertainty source (i.e. GCM structure), the

second largest contributor varied between emission scenario and bias correction for the individual seasons. Future work should investigate if the selection of the snow model has a stronger impact in other regions or catchments of different size (i.e. small headwater sub-catchments). Model parameter uncertainty was not accounted for in this study but should be investigated as part of future work, which could help to understand and potentially improve misrepresentations in the historic runoff regime.

5 The use of other hydrological indicators (i.e. low and high flow) should also be explored as the effect of the individual components of the model chain might differ for such alternative metrics.

Acknowledgements

We would like to thank a number of authorities who provided the various data sets used in this study. The National Institute of Water and Atmospheric Research (NIWA) for climate observations and the RCM simulations. Streamflow records were

10 kindly provided by NIWA, Contact Energy Ltd. and the Otago Regional Council. The digital elevation model, soil data and land use information were obtained from Landcare Research New Zealand. This study was funded by a University of Otago Doctoral Scholarship and Publishing Bursary.

Competing interests

The authors declare that they have no conflict of interest.

15 References

- Ackerley, D., Dean, S., Sood, A., Mullan, A. B.: Regional climate modelling in New Zealand: Comparison to gridded and satellite observations, *Weather and Climate*, 32(1), 3-22, 2012.
- Anderson, E. A.: National Weather Service river forecast system - snow accumulation and ablation model, National Oceanographic and Atmospheric Administration (NOAA), Technical Memorandum, NWS-HYDRO-17, U.S. Department of
- 20 Commerce, Silver Spring, MD. 238 pp, 1973.
- Bennett, K. E., Werner, A. T., Schnorbus, M.: Uncertainties in hydrologic and climate change impact analyses in headwater basins of British Columbia, *J. Climate*, 25, 5711-5730, doi:10.1175/JCLI-D-11-00417.1, 2012.
- Boé, J., Terray, L., Habets, F., Martin, E.: Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies, *Int. J. Climatol.*, 27, 1643–1655, doi:10.1002/joc.1602, 2007.
- 25 Caruso, B. S., King, R., Newton, S., Zammit, C.: Simulation of climate change effects on hydropower operations in mountain headwater lakes, New Zealand, *River Res. Appl.*, 33, 147-161, doi:10.1002/rra.3056, 2017.
- Chang, H., Jung, I.: Spatial and temporal changes in runoff caused by climate change in a complex large river basin in Oregon, *J. Hydrol.*, 388, 186–207, doi:10.1016/j.jhydrol.2010.04.040, 2010.

- Chen, J., Brissette, F. P., Chaumont, D., Braun, M.: Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America, *Water Resour. Res.*, 49, 4187–4205, doi:10.1002/wrcr.20331, 2013.
- Clark, M. P., Rupp, D. E., Woods, R. A., Zheng, X., Ibbitt, R. P., Slater, A. G., Uddstrom, M. J.: Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model, *Adv. Water Resour.*, 31(10), 1309-1324, doi:10.1016/j.advwatres.2008.06.005, 2008.
- 5 Cullen, N. J., Conway, J. P.: A 22-month record of surface meteorology and energy balance from the ablation zone of Brewster Glacier, New Zealand, *J. Glaciol.*, 61, 931-946, doi:10.3189/2015JG15J004, 2015.
- De Elia, R., Cote, H.: Climate and climate change sensitivity to model configuration in the Canadian RCM over North America, *Meteorologische Zeitschrift*, 19(4), 325-339, 2010.
- 10 Dobler, C., Hagemann, S., Wilby, R. L., Stötter, J.: Quantifying different sources of uncertainty in hydrological projections in an Alpine watershed, *Hydrol. Earth Syst. Sc.*, 16, 4343–4360, doi:10.5194/hess-16-4343-2012, 2012.
- [Freudiger, D., Kohn, I., Seibert, J., Stahl, K., Weiler, M.: Snow redistribution for the hydrological modeling of alpine catchments, *WIREs Water*, 4, e1232, doi: 10.1002/wat2.1232, 2017.](#)
- Frueh, B., Schipper, J. W., Pfeiffer, A., Wirth, V. A.: A pragmatic approach for downscaling precipitation in alpine-scale complex terrain, *Meteorologische Zeitschrift*, 15(6), 631-646, 2006.
- 15 Gawith, D., Kingston, D. G., McMillan, H.: The effects of climate change on runoff in the Lindis and Matukituki catchments, Otago, New Zealand, *Journal of Hydrology New Zealand*, 51, 121-135, 2012.
- Graham, L. P., Hagemann, S., Jaun, S., Beniston, M.: On interpreting hydrological change from Regional Climate Models, *Climatic Change*, 81, 97–122, 2007.
- 20 Gudmundsson, L., Tallaksen, L. M., Stahl, K., Clark, D. B., Dumont, E., Hagemann, S., Bertrand, N., Gerten, D., Heinke, J., Hanasaki, N., Voss, F., Koirala, S.: Comparing large-scale hydrological model simulations to observed runoff percentiles in Europe, *J. Hydrometeorol.*, 13, 604–620, doi:10.1175/JHM-D-11-083.1, 2012.
- Gutjahr, O., Heinemann, G.: Comparing precipitation bias correction methods for high resolution regional climate simulations using Cosmo-Clm: Effects on extreme values and climate change signal, *Theor. Appl. Climatol.*, 114(3-4), 511-529, 2013.
- 25 Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voß, F., Arnell, N. W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacke, T., Viterbo, P., Weedon, G. P., Yeh, P.: Multimodel estimate of the global terrestrial water balance: Setup and first results, *J. Hydrometeorol.*, 12, 869–884, doi:10.1175/2011JHM1324.1, 2011.
- Hagemann, S., Chen, C., Haerter, J. O., Heinke, J., Gerten, D., Piani, C.: Impact of a statistical bias correction on the projected hydrological changes obtained from three GCMs and two hydrology models, *J. Hydrometeorol.*, 12(4), 556-578, 2011.
- 30 Henderson, R. D., Thompson, S. M.: 1999. Extreme rainfalls in the Southern Alps of New Zealand, *Journal of Hydrology New Zealand*, 38(2), 309-330, 2011.
- Hendrikx, J., Hreinsson, E.Ö., Clark, M.P., Mullan, A.B: The potential impact of climate change on

- seasonal snow in New Zealand: Part I—an analysis using 12 GCMs, *Theor. Appl. Climatol.*, 110, 607-618, doi:10.1007/s00704-012-0711-1, 2012.
- Hidalgo, H. G., Das, T., Dettinger, M. D., Cayan, D. R., Pierce, D.W., Barnett, T. P., Bala, G., Mirin, A., Wood, A. W., Bonfils, C., Santer, B. D., Nozawa, T.: Detection and attribution of streamflow timing changes to climate change in the western United States, *J. Climate*, 22, 3838-3855, doi:10.1175/2009JCLI2470.1, 2009.
- Hughes, D. A., Kingston, D. G., Todd, M. C.: Uncertainty in water resources availability in the Okavango River basin as a result of climate change, *Hydrol. Earth Syst. Sc.*, 15, 931–941, 2011.
- IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A., Nauels, Y., Xia, V., Bex and P.M. Midgley (eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp, 2013.
- Jiménez Cisneros, B. E., Oki, T., Arnell, N. W., Benito, G., Cogley, J. G., Döll, P., Jiang, T., Mwakalila, S. S.: Freshwater resources, In: *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 229-269, 2014.
- Jobst, A. M.: The potential impacts of climate change on the hydro-climate of the Clutha/Mata-Au catchment. Ph. D. thesis, Department of Geography, University of Otago, Dunedin, New Zealand, 334 pp, 2017.
- Jobst, A. M., Kingston, D. G., Cullen, N. J., Sirguey, P.: Combining thin-plate spline interpolation with a lapse rate model to produce daily air temperature estimates in a data-sparse alpine catchment, *Int. J. Climatol.*, 37(1), 214–229, doi:10.1002/joc.4699, 2017.
- Jung, I., Moradkhani, H., Chang, H.: Uncertainty assessment of climate change impacts for hydrologically distinct river basins, *J. Hydrol.*, 466–467, 73-87, doi:10.1016/j.jhydrol.2012.08.002, 2012.
- Kay, A. L., Davies, H. N.: Calculating potential evaporation from climate model data: a source of uncertainty for hydrological climate change impacts, *J. Hydrol.*, 358(3-4), 221-239, doi:10.1016/j.jhydrol.2008.06.005, 2008.
- Kennedy, J., Eberhart, R.: Particle swarm optimization, Paper presented at the IEEE International Conference on Neural Networks - Conference Proceedings, 1995.
- Kingston, D. G., Taylor, R. G.: Sources of uncertainty in climate change impacts on river discharge and groundwater in a headwater catchment of the Upper Nile Basin, Uganda, *Hydrol. Earth Syst. Sc.*, 14, 1297–1308, 2010.
- Koedyk, L. P., Kingston, D. G.: Potential evapotranspiration method influence on climate change impacts on river flow: a mid-latitude case study, *Hydrology Research*, doi:10.2166/nh.2016.152, 2016.
- Kundzewicz, Z. W., Mata, L. J., Arnell, N. W., Döll, P., Jimenez, B., Miller, K., Oki, T., Sen, Z., Shiklomanov, I.: The implications of projected climate change for freshwater resources and their management, *Hydrolog. Sci. J.* 53(1), 3-10, 2008.

- Lafon, T., Dadson, S., Buys, G., Prudhomme, C.: Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods, *Int. J. Climatol.*, 33, 1367–1381, doi:10.1002/joc.3518, 2013.
- Laghari, A. N., Vanham, D., Rauch, W.: To what extent does climate change result in a shift in Alpine hydrology? A case study in the Austrian Alps, *Hydrolog. Sci. J.*, 57, 103-117, doi:10.1080/02626667.2011.637040, 2012.
- 5 Lenderink, G., Buishand, A., Van Deursen, W.: Estimates of future discharges of the river Rhine using two scenario methodologies: Direct versus delta approach, *Hydrol. Earth Syst. Sc.*, 11(3), 1145-1159, 2007.
- Leung, L.R., Qian, Y., Bian, X., Washington, W.M., Han, J., Roads, J. O.: Mid-century ensemble regional climate change scenarios for the western United States, *Climatic Change*, 62(1-3), 75-113, 2004.
- Ludwig, R., May, I., Turcotte, R., Vescovi, L., Braun, M., Cyr, J. F., Fortin, L. G., Chaumont, D., Biner, S., Chartier, I., Caya,
10 D., Mauser, W.: The role of hydrological model complexity and uncertainty in climate change impact assessment, *Advances in Geosciences*, 21, 63-71, 2009.
- Lutz, A. F., Immerzeel, W. W., Kraaijenbrink, P. D. A., Shrestha, A.B, Bierkens, M. F. P.: Climate Change Impacts on the Upper Indus Hydrology: Sources, Shifts and Extremes, *PLOS ONE*, 11, e0165630, doi:10.1371/journal.pone.0165630, 2016.
- Mandal, S., Simonovic, S. P.: Quantification of uncertainty in the assessment of future streamflow under changing climate
15 conditions, *Hydrol. Process.*, 31, 2076–2094, doi:10.1002/hyp.11174, 2017.
- Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, *Rev. Geophys.*, 48, doi:10.1029/2009rg000314, 2010.
- 20 Marke, T., Mauser, W., Pfeiffer, A., Zängl, G.: A pragmatic approach for the downscaling and bias correction of regional climate simulations: evaluation in hydrological modelling, *Geoscientific Model Development* 4(3), 759-770, doi:10.5194/gmd-4-759-2011, 2011.
- Ministry for the Environment: Climate change effects and impacts assessment. A guidance manual for local government in New Zealand, 2nd Edition, Prepared by Mullan, B.; Wratt, D.; Dean, S.; Hollis, M. (NIWA); Allan, S.; Williams, T. (MWH
25 NZ Ltd) and Kenny, G. (Earthwise Consulting Ltd), in consultation with Ministry for the Environment, NIWA Client Report No. WLG2007/62., 156 pp, 2008.
- Ministry for the Environment: Climate Change Projections for New Zealand: Atmosphere projections based on simulations from the IPCC Fifth Assessment, Prepared by Mullan, B.; Sood, A.; Stuart, S. (NIWA), Wellington, 127 pp, 2016.
- Mpelasoka, F. S., Chiew, F. H. S.: Influence of rainfall scenario construction methods on runoff projections, *J. Hydrometeorol.*,
30 10(5), 1168-83, 2009.
- Muerth, M. J., St-Denis, G. B., Ludwig, R., Caya, D.: Evaluation of different sources of uncertainty in climate change impact research using a hydro-climatic model ensemble, R. Seppelt, A.A. Voinov, S. Lange, D. Bankamp (Eds.): *International Environmental Modelling and Software Society (iEMSs) 2012 International Congress on Environmental Modelling and*

- Software Managing Resources of a Limited Planet: Pathways and Visions under Uncertainty, Sixth Biennial Meeting, Leipzig, Germany, 2012.
- Muerth, M. J., St-Denis, G. B., Ricard, S., Velázquez, J. A., Schmid, J., Minville, M., Caya, D.: On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff, *Hydrol. Earth Syst. Sc.*, 17(3), 1189-204, 2013.
- 5 Murray, D. L.: Regional hydrology of the Clutha River, *Journal of Hydrology New Zealand*, 14(2), 83-98, 1975.
- Palazzi, E., von Hardenberg, J., Terzago, S., Provenzale, A.: Precipitation in the Karakoram-Himalaya: a CMIP5 view, *Clim. Dynam.*, 45, 21–45, doi:10.1007/s00382-014-2341-z, 2014.
- Pall, P., Aina, T., Stone, D. A., Stott, P. A., Nozawa, T., Hilberts, A. G. J., Lohmann, D., Allen, M. R.: Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000, *Nature*, 470(7334), 382-385, doi:10.1038/nature09762, 2011.
- 10 Piao, S., Ciais, P., Huang, Y., Shen, Z., Peng, S., Li, J., Zhou, L., Liu, H., Ma, Y., Ding, Y., Friedlingstein, P., Liu, C., Tan, K., Yu, Y., Zhang, T., Fang, J.: The impacts of climate change on water resources and agriculture in China, *Nature*, 467, 43-51, doi:10.1038/nature09364, 2010.
- Poyck, S., Hendrikx, J., McMillan, H., Hreinsson, E. Ö., Woods, R.: Combined snow and streamflow modelling to estimate impacts of climate change on water resources in the Clutha River, New Zealand, *Journal of Hydrology New Zealand*, 50(2), 293-311, doi:10.1007/s00704-012-0711-1, 2011.
- 15 Prowse, T. D., Owens, I. F.: Energy balance over melting snow, Craigieburn Range, New Zealand, *Journal of Hydrology New Zealand*, 21(2), 133-147, 1982.
- Prudhomme, C., Davies, H.: Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 2: Future climate, *Climatic Change*, 93(1-2), 197-222, 2009.
- 20 Rosenzweig, C., Casassa, G., Karoly, D. J., Imeson, A., Liu, C., Menzel, A., Rawlins, S., Root, T. L., Seguin, B., Tryjanowski P.: Assessment of observed changes and responses in natural and managed systems. In: *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Parry, M.L., O.F. Canziani, J.P. Palutikof, P.J. van der Linden, and C.E. Hanson (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA, 79-131, 2007.
- 25 Schulla, J.: Model description WaSiM, Technical report, 324 pp, Available at <http://wasim.ch>, 2012.
- Shrestha, R. R., Schnorbus, M. A., Werner, A. T., Berland, A. J.: Modelling spatial and temporal variability of hydrologic impacts of climate change in the Fraser River basin, British Columbia, Canada, *Hydrol. Process.*, 26, 1841-1861, doi:10.1002/hyp.9283, 2012.
- 30 Sims, C., Orwin, J. F.: Snowmelt generation on a hydrologically sensitive mountain range, Pisa Range, Central Otago, New Zealand, *Journal of Hydrology New Zealand*, 50(2), 383-405, doi:10.1002/joc.2216, 2011.
- Sinclair, M. R., Wratt, D. S., Henderson, R. D., Gray, W. R.: Factors affecting the distribution and spillover of precipitation in the Southern Alps of New Zealand—A Case Study, *J. Appl. Meteorol.*, 36(5), 428-442, doi:10.1175/1520-0450(1997)036<0428:FATDAS>2.0.CO;2, 1997.

- Srinivasan, M.S., Schmidt, J., Poyck, S., Hreinsson, E.: Irrigation reliability under climate change scenarios: A modeling investigation in a river-based irrigation scheme in New Zealand, *J. Am. Water Resour. As.* 47(6), 1261-1274, doi:10.1111/j.1752-1688.2011.00568.x, 2011.
- Sturman, A. P., McGowan, A. H., Spronken-Smith, A. R.: Mesoscale and local climates in New Zealand, *Prog. Phys. Geog.*, 23(4), 611-635, doi:10.1177/030913339902300407, 1999.
- 5 Tait, A.B., Sood, A., Mullan, B., Stuart, S., Bodeker, G., Kremser, S., Lewis, J.: Updated climate change projections for New Zealand for use in impact studies. Synthesis report RA1. Climate Changes, Impacts and Implications (CCII) for New Zealand to 2100. MBIE contract C01X1225. 18 pp, 2016.
- Teng, J., Vaze, J., Chiew, F. H. S., Wang, B., Perraud, J. M.: Estimating the relative uncertainties sourced from GCMs and hydrological models in modeling climate change impact on runoff, *J. Hydrometeorol.*, 13, 122-139, 2012.
- 10 Teutschbein, C., Seibert, J.: Regional Climate Models for hydrological impact studies at the catchment scale: A review of recent modeling strategies, *Geography Compass* 4(7), 834-60, 2010.
- Thompson, J. R., Green, A. J., Kingston, D. G., Gosling, S. N.: Assessment of uncertainty in river flow projections for the Mekong River using multiple GCMs and hydrological models, *J. Hydrol.*, 486, 1-30, 2013.
- 15 Thompson, J. R., Green, A. J., Kingston, D. G.: Potential evapotranspiration-related uncertainty in climate change impacts on river flow: An assessment for the Mekong River basin, *J. Hydrol.*, 510, 259-279, doi:10.1016/j.jhydrol.2013.12.010, 2014.
- Troin, M., Poulin, A., Baraer, M., Brissette, F.: Comparing snow models under current and future climates: Uncertainties and implications for hydrological impact studies, *J. Hydrol.*, 540, 588-602, doi:10.1016/j.jhydrol.2016.06.055, 2016.
- Van Vuuren, D. P., Edmonds, J., Kainuma, M., et al.: The Representative Concentration Pathways: an overview, *Climatic Change* 109, 5, doi:10.1007/s10584-011-0148-z, 2011.
- 20 Velázquez, J. A., Schmid, J., Ricard, S., Muerth, M. J., Gauvin St-Denis, B., Minville, M., Chaumont, D., Caya, D., Ludwig, R., Turcotte, R.: An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources, *Hydrol. Earth Syst. Sc.*, 17(2), 565-578, doi:10.5194/hess-17-565-2013, 2013.
- Vormoor, K., Lawrence, D., Heistermann, M., Bronstert, A.: Climate change impacts on the seasonality and generation processes of floods – projections and uncertainties for catchments with mixed snowmelt/rainfall regimes, *Hydrol. Earth Syst. Sc.*, 19, 913-931, doi:10.5194/hess-19-913-2015, 2015.
- 25 Wilby, R. L., Harris, I.: A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK, *Water Resour. Res.*, 42, doi:10.1029/2005wr004065, 2006.
- [Willsman, A.: Annual glacier volumes in New Zealand 1995-2010. NIWA Client Report, No CHC2011-050, 2011.](#)
- 30 Zammit, C., Woods, R.: Projected climate and river flow for the Waimakariri catchment for 2040s and 2090s. NIWA Client Report, No. CHCH2011-025, 2011.
- Zemansky, G., Hong, Y. S. T., Rose, J., Song, S. H., Thomas, J.: Assessing the effects of climate change on water resources: The Waimea Plains, *Journal of Hydrology New Zealand*, 51(1), 45-62, 2012.

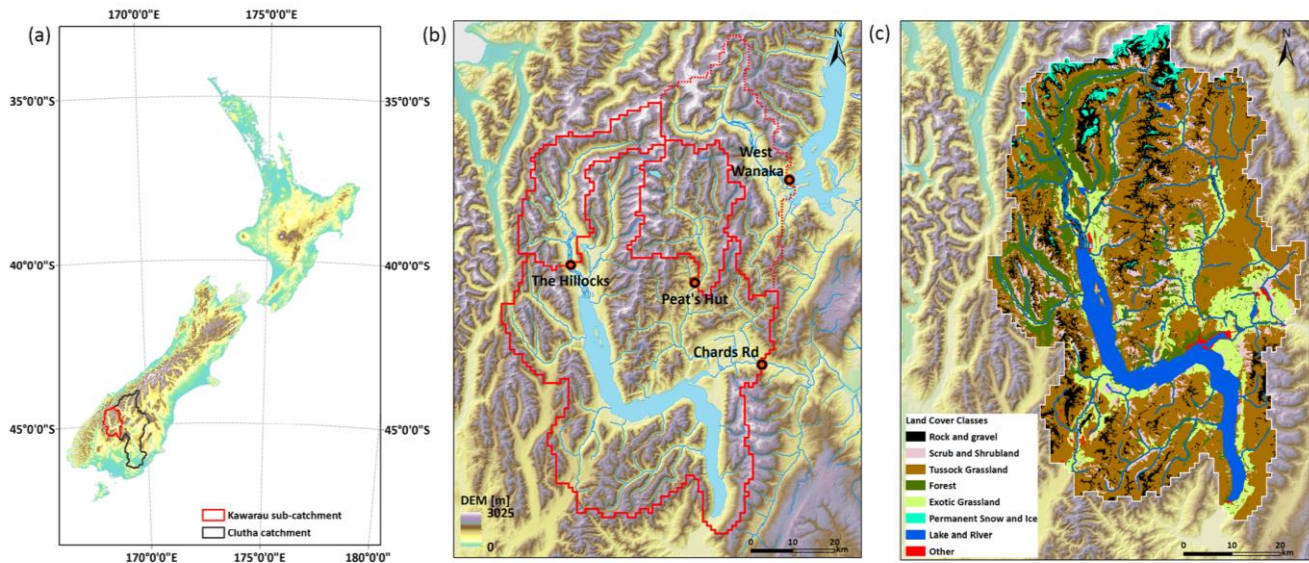
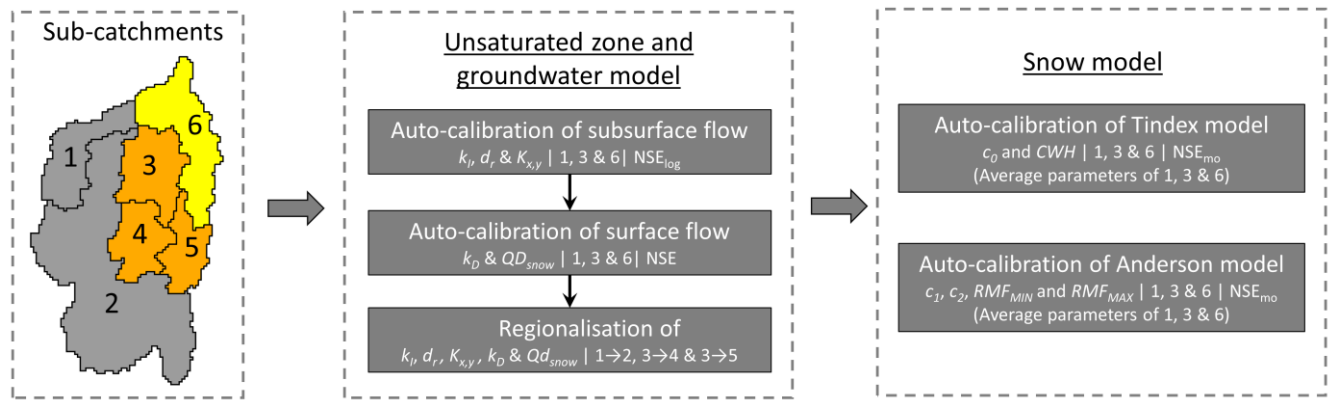


Figure 1 Maps showing (a) New Zealand with the Clutha catchment located in the lower South Island, (b) the sub-catchments of the gauges recording discharge used in this study (note that the West Wanaka sub-catchment is outside the Kawarau catchment) and (c) a land cover classification of the Kawarau sub-catchment based on New Zealand's Land Cover Database (LCDB v3.0) that was published by Landcare Research in 2012.

5



10 **Figure AA-The2** The calibration workflow that was used to calibrate WaSiM. For each calibration step (either manual or automatic (Auto-cal)) information is provided as follows: “calibrated parameters | calibrated sub-catchments | performance criterium used”. The parameters controlling subsurface flow encompass the drainage density for interflow (d_r), the storage coefficient of interflow (k_f) and the groundwater conductivity in x or y-direction. Surface flow is controlled by the storage coefficient of surface runoff (k_D) and the fraction of snowmelt directly becoming surface runoff (OD_{snow}). The parameters of the snow models are shown in Table 1.

15

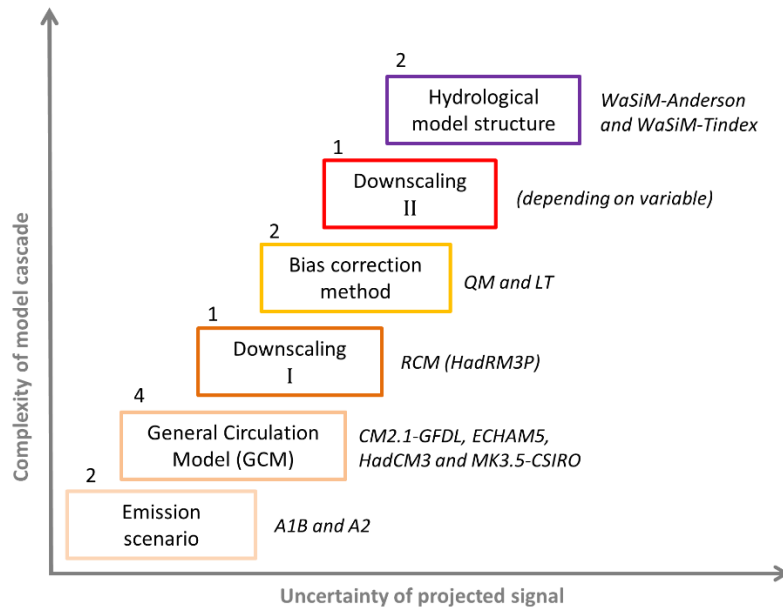


Figure 32 The model cascade is depicted with the individual members of each component listed on the right. Permuting the members of all components results in a total of 32 hydrological projections.

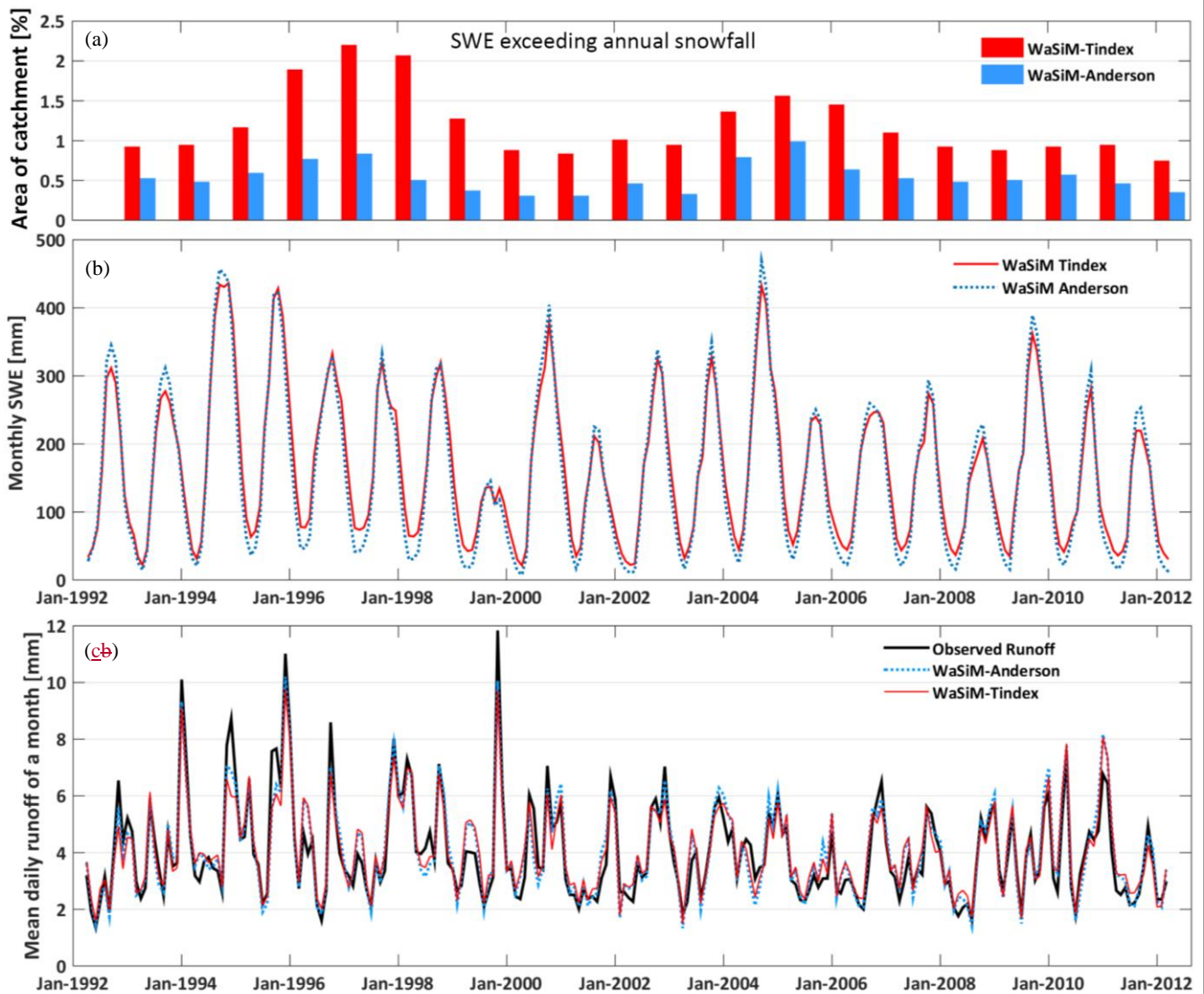
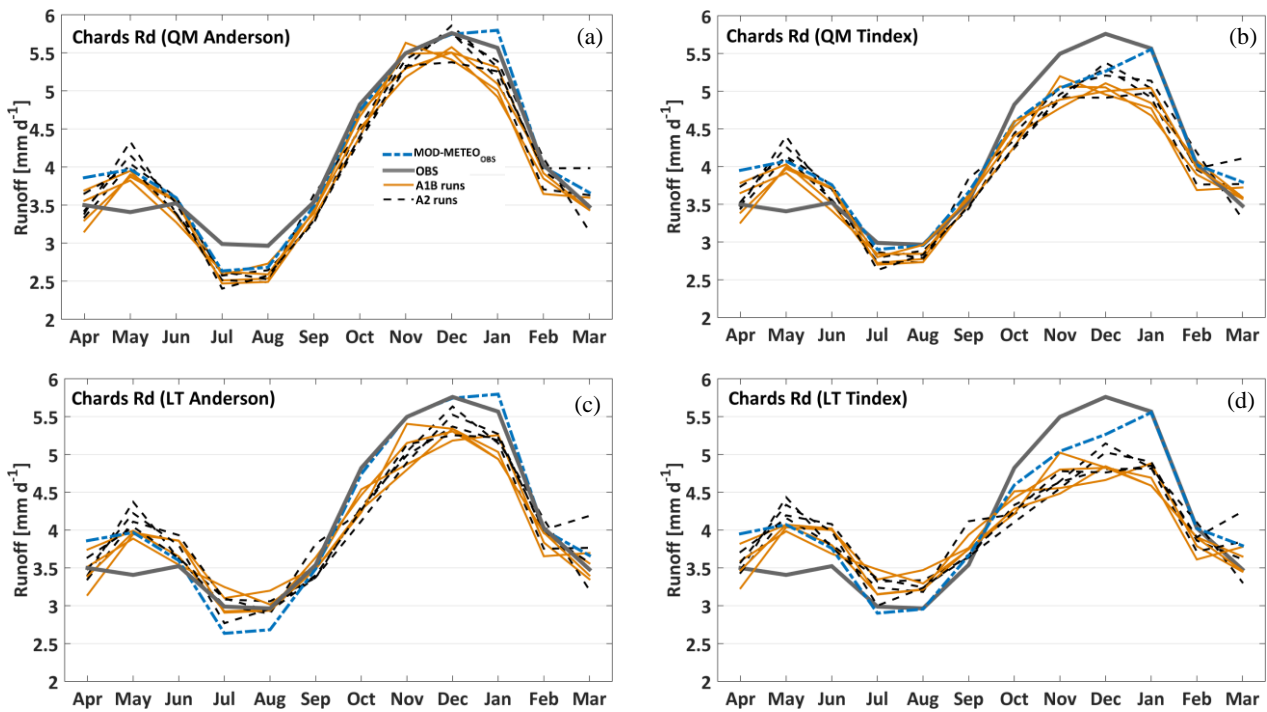
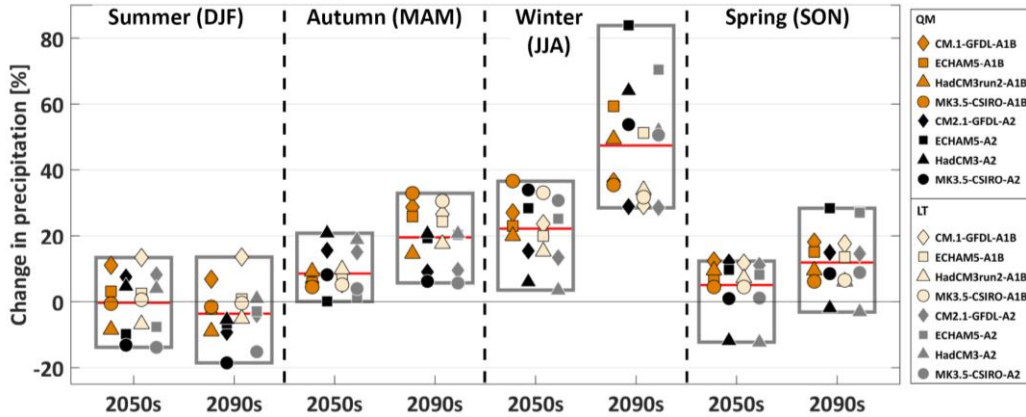


Figure 43 (a) relative area of the Kawarau catchment where SWE exceeds annual snowfall for a hydrological year (e.g. 1.4.1992 – 31.3.1993) as an indicator for unrealistic build-up of snow (i.e. snow towers). (b) Modelled mean monthly snow water equivalent (SWE) based on the Tindex and Anderson simulations. (c) Observed and modelled (WaSiM-Anderson and WaSiM-Tindex) monthly runoff at Chards Rd for the reference period (1992-2012).



5 **Figure 54** RCM driven mean daily runoff simulations for the 1/4/1992–30/3/2011 period at Chards Rd. Simulations are compared with the observed regime (grey line) and the modelled (WaSiM forced with observed meteorology) regime (blue line) for the four subensembles: QM-Anderson (a), QM-Tindex (b), LT-Anderson (c) and LT-Tindex (d).



10 **Figure 65** The uncertainty range of the precipitation signal (domain average) is shown for the entire ensemble and the four seasons. For each box the uncertainty range is broken down into the two emission scenarios and the two bias correction methods. The red line represents the median of all 16 members.

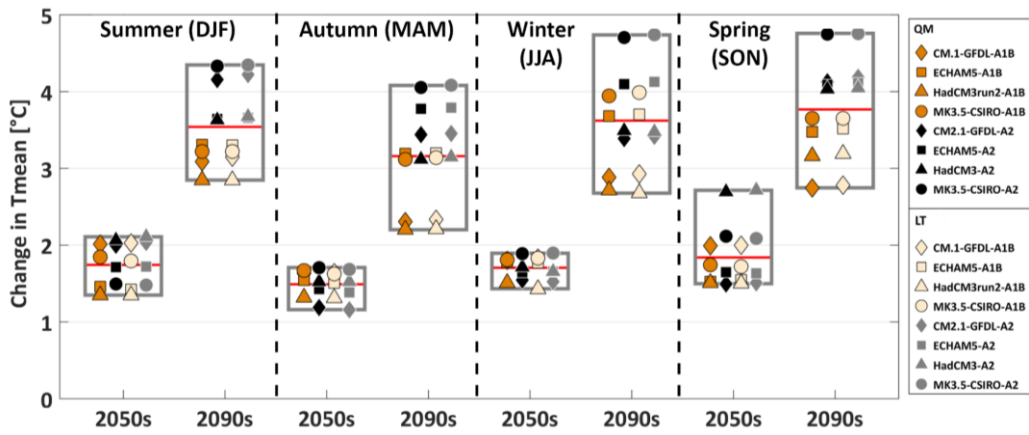


Figure 76 The uncertainty range of the Tmean signal is shown for the entire ensemble and the four seasons. For each box the uncertainty range is broken down into the two emission scenarios and the two bias correction methods. The red line represents the median of all 16 members within a box.

5

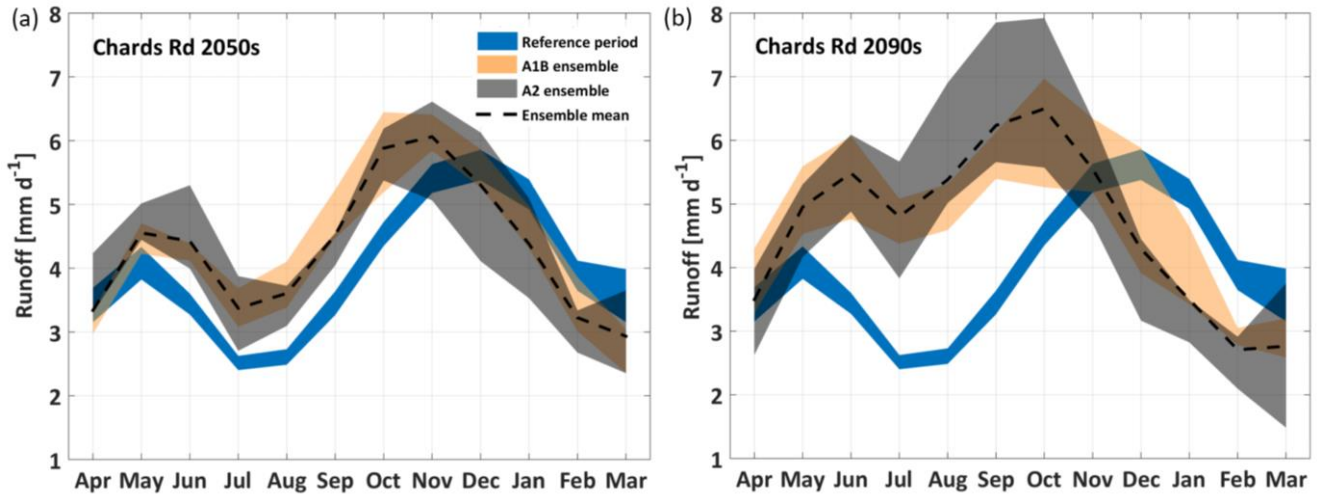


Figure 87 Mean daily modelled runoff (orange envelope = A1B and grey envelope = A2) at Chards Rd during the 2050s (a) and 2090s (b) is compared with the historic simulations (blue). All simulations are based on the QM-Anderson subensemble.

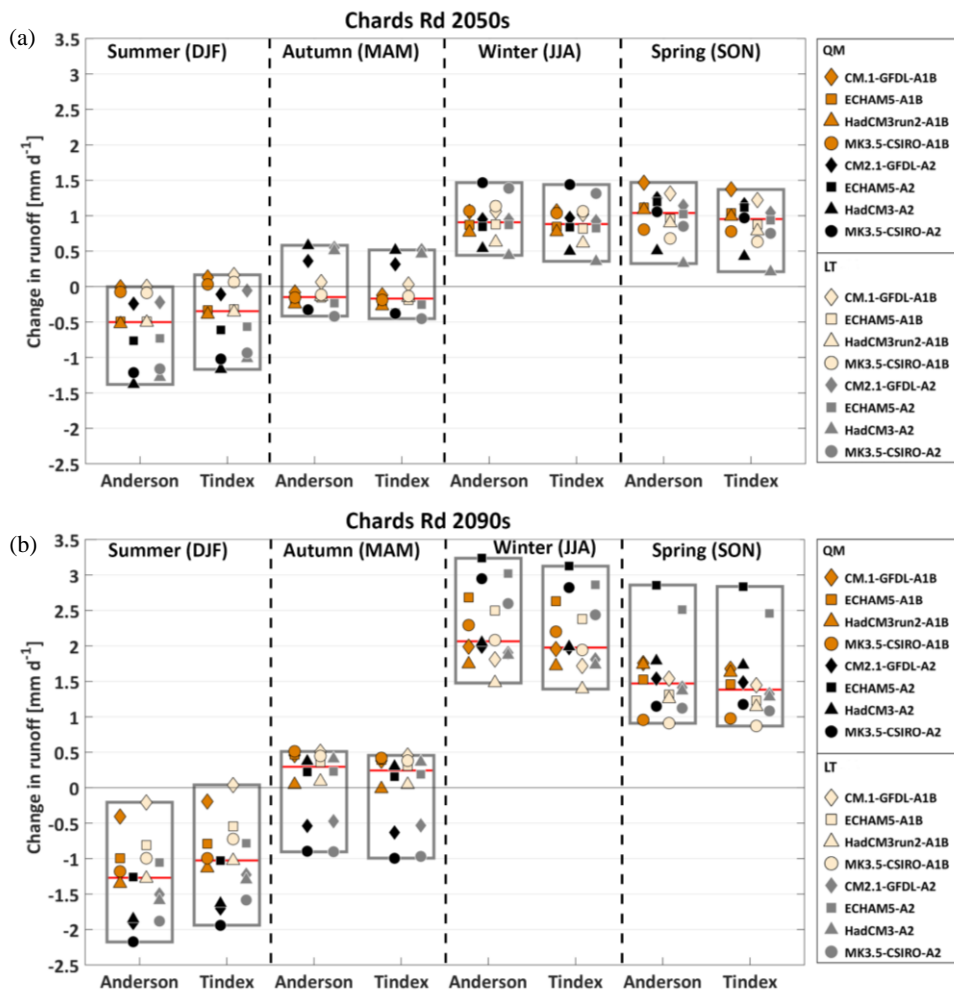


Figure 98 The uncertainty range of the seasonal runoff signal at Chards Rd is shown for (a) the 2050s and (b) the 2090s. For each season an uncertainty box is shown for WaSiM-Anderson and WaSiM-Tindex simulations, respectively. In each box (red line = median) the uncertainty range is broken down into the two emission scenarios and the two bias correction methods (LT and QM).

5

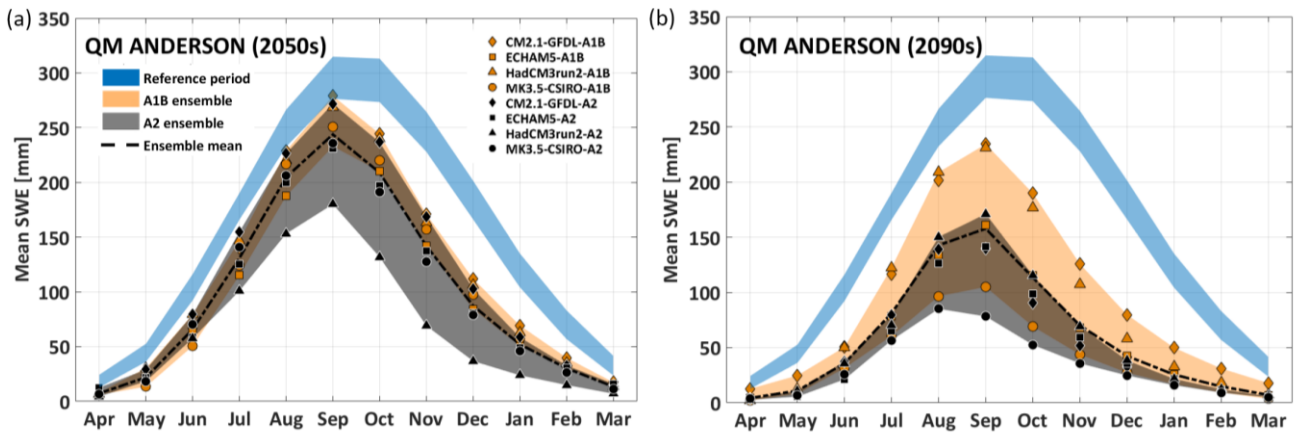


Figure 109 The monthly SWE volume simulated by the QM-Anderson subensemble (km^3) is depicted for (a) the 2050s and (b) the 2090s period. The individual members are augmented by the A1B envelope, the A2 envelope and the ensemble mean (dashed line). The blue envelope represents the range of SWE volume simulations during the baseline period.

5

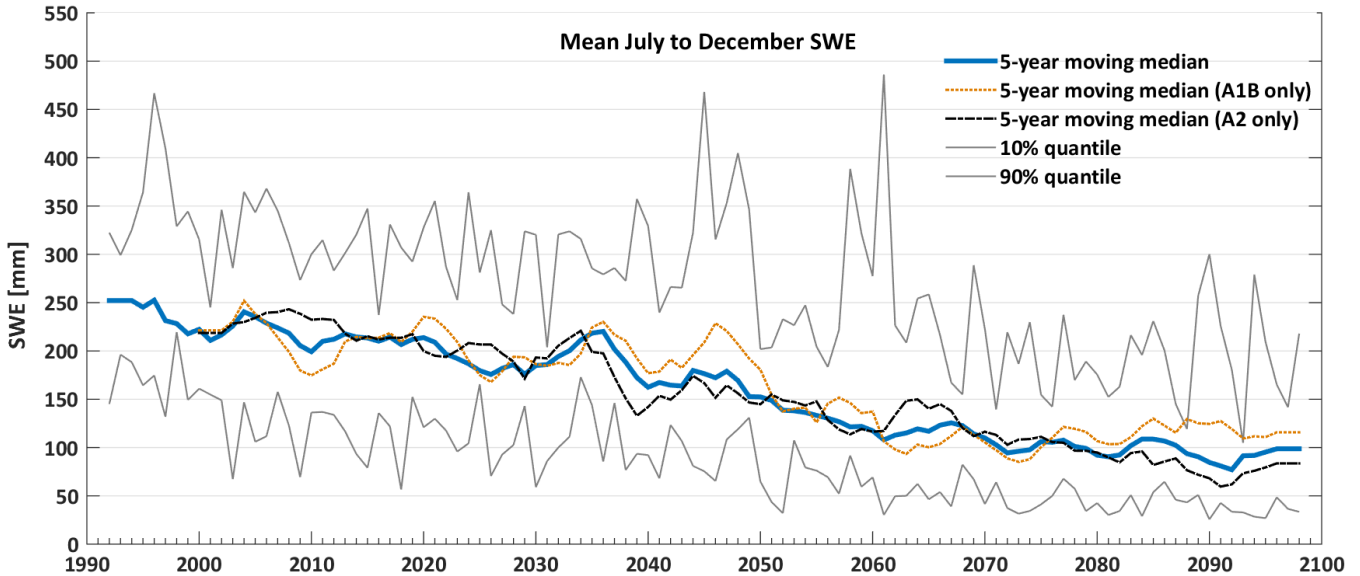


Figure 11BB The mean SWE of the six months with the historically highest SWE (July to December) is shown for the 16 transient QM simulations (including QM-Anderson and QM-Tindex).

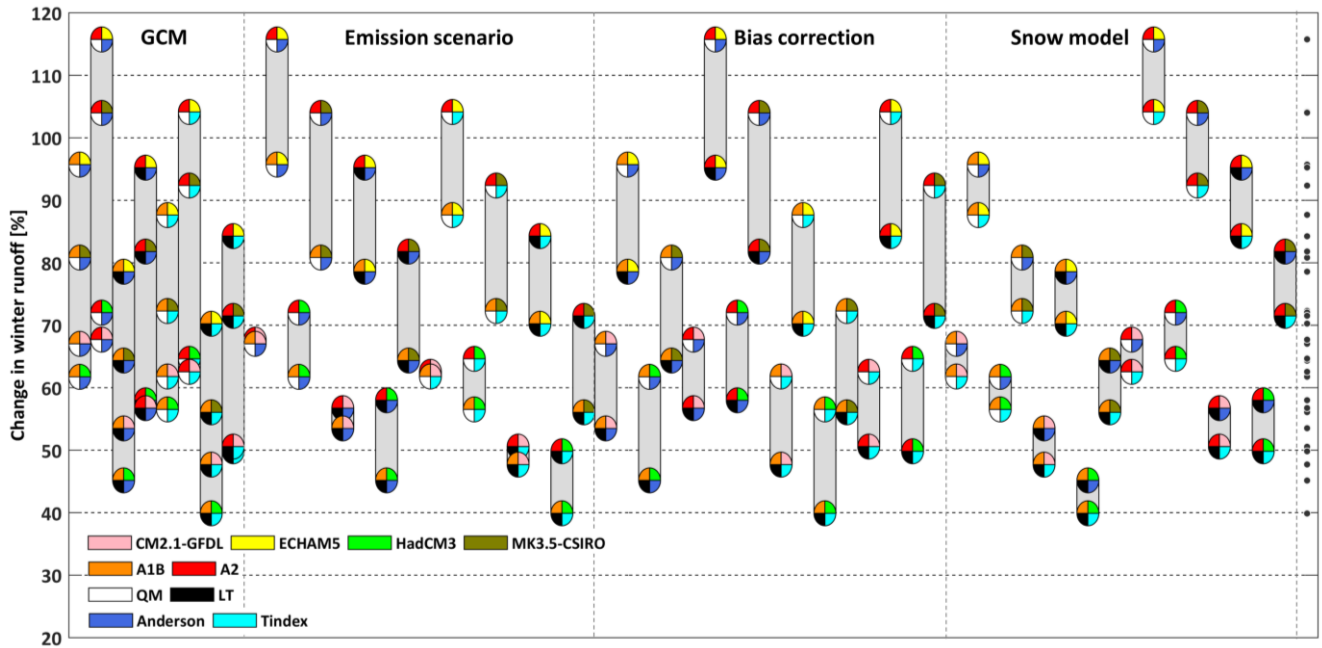


Figure 12: Uncertainty contributions are isolated for each of the four model components GCM, emission scenario, bias correction method and snow model (the approach is described in more detail in section 3.4). The number of points in a bar corresponds to the number of members of the selected model component (e.g. four GCMs), while the number of bars is equal to the possible permutations based on the remaining three model components ($2 \cdot 2 \cdot 2 = 8$).

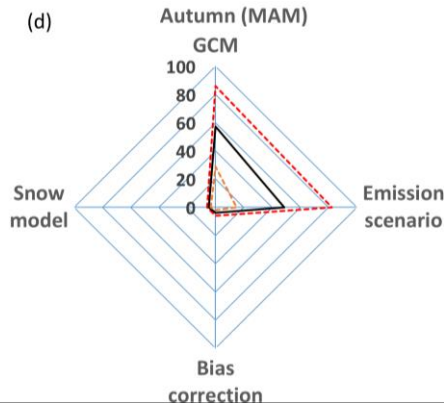
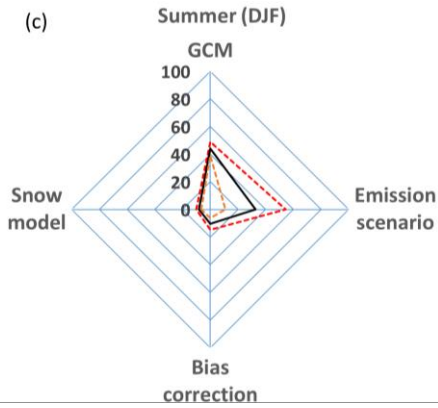
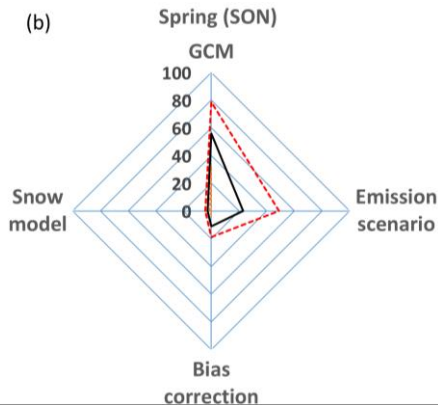
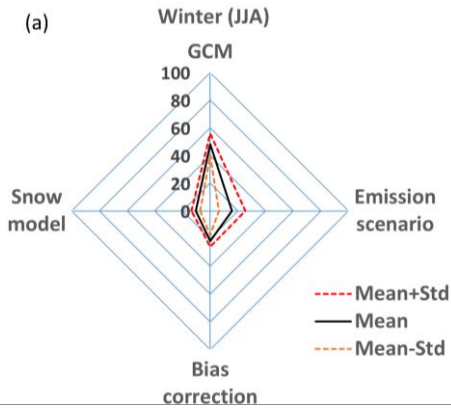


Figure 13 Relative contribution of the four uncertainty sources to the overall uncertainty of the seasonal runoff signal for (b) winter, (c) spring, (d) summer and (e) autumn. Based on the permutations (as exemplified for winter in Figure 10) the radar charts show the mean contributions [%] (mean length of bars shown in Figure 10) of the four model components to the overall uncertainty as well as the standard deviations [%] (standard deviation of bars shown in Figure 10).

10

15

Table 1 Parameters of the two snow-melt routines as determined by the calibration routine. T_0 and CWH are mutual parameters of both models. The snow accumulation parameters are based on the findings of Auer (1974).

Anderson and Tindex		
Threshold temperature for snowmelt (T_0)	0.00	[°C]
Water holding capacity of snow (CWH)	0.29	[-]
Tindex		
Temperature dependent DDF (c_0)	1.91	[mm °C ⁻¹ d ⁻¹]
Anderson		
Temperature dependent DDF (c_1)	0.63	[mm °C ⁻¹ d ⁻¹]
Wind dependent DDF (c_2)	0.08	[mm (°C m/s d) ⁻¹]
Minimum radiation melt factor (RMF_{MAX})	3.13	[mm °C ⁻¹ d ⁻¹]
Maximum radiation melt factor (RMF_{MIN})	0.36	[mm °C ⁻¹ d ⁻¹]
Snow accumulation		
Temperature, at which 50% of precipitation are falling as snow (T_{RS})	3.00	[°C]
Half of the temperature-transition range from snow to rain (T_{trans})	3.00	[K]

5 **Table 2 Nash Sutcliffe values based on daily (NSE), logarithmic (NSElog) and monthly streamflow data during the calibration (Cal) and validation (Val) periods calculated for the WaSiM-Anderson and WaSiM-Tindex (in brackets) simulations. Note the different validation periods for The Hillocks and Peat's Hut due to shorter records.**

River	Gauge	Cal (1.4.2008-31.3.2012)			Val (1.4.1992-31.3.2008)		
		NSE	NSE _{log}	NSE _{mo}	NSE	NSE _{log}	NSE _{mo}
Dart	The Hillocks (1996-2012)	0.77 (0.77)	0.77 (0.78)	0.92 (0.92)	0.64 (0.65)	0.64 (0.68)	0.78 (0.79)
Shotover	Peat's Hut (1996-2012)	0.64 (0.65)	0.67 (0.70)	0.81 (0.82)	0.60 (0.62)	0.65 (0.70)	0.76 (0.79)
Kawarau	Chards Rd	0.87 (0.88)	0.88 (0.87)	0.89 (0.90)	0.87 (0.85)	0.86 (0.86)	0.89 (0.87)
Matukituki	West Wanaka	0.67 (0.67)	0.64 (0.65)	0.80 (0.80)	0.62 (0.62)	0.72 (0.72)	0.83 (0.82)

10 **Table 3 The historic (1/4/1992 – 30/3/2011) water balance terms of the MOD-METEO_{OBS} run compared with the corresponding (depending on snow model) RCM ensemble means (QM and LT). The seasonal and annual SWE volumes are also shown.**

Anderson	P	ET	Q	SWE [mm]				
	[mm]	[mm]	[mm]	DJF	MAM	JJA	SON	YEAR
MOD-METEO _{OBS}	2007	471	1512	117	33	183	286	154
QM	1926	473	1425	123	31	176	277	152
LT	1931	466	1443	117	33	150	240	134
Tindex								
MOD-METEO _{OBS}	2007	473	1509	143	53	181	275	163
QM	1926	475	1422	143	44	172	262	154
LT	1931	468	1441	126	42	139	216	130