We would like to thank the second reviewer for the constructive comments. As already explained under our response to reviewer one we have clarified the geographical domain of this modelling study throughout the document and believe that the geographical focus of this study is now more obvious. We have also provided more detail about the snow models used, and presented more results from the validation of the hydrological model including some more performance criteria. Please see below for our individual responses.

H. McMillan (Referee) hmcmillan@sdsu.edu Received and published: 7 January 2018 Review of "Intercomparison of different uncertainty sources in hydrological climate change projections for an alpine catchment (Clutha River, New Zealand)" by Jobst et al. This paper is an exploration of the potential impacts of climate change on precipitation, snowpack and river flow in a sub-catchment of the Clutha River in New Zealand. The paper is well written and comprehensive, and I recommend it for publication after the

minor revisions outlined below.

1. The start of the paper makes much of the large size and representativeness of the Clutha catchment. Given that the study is actually only carried out on a sub catchment (less than one quarter area) of the Clutha, which does not include any of the drier Otago climate described, I suggest this section be revised for relevance.

Response: *See answers to comment 1 of reviewer 1.*

2. p5 L15 The authors should define in the text what they mean by the "observed regime" (i.e. monthly flow values averaged across all years) so that the readers are clear what is being evaluated. Similarly the meaning of "summer climate change signal" should be defined.

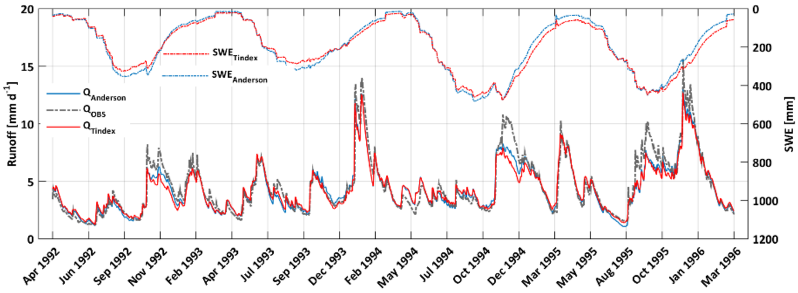
Response: *Sentence has been changed to: "The regimes (i.e. monthly flow values averaged across all years) of the eight RCM driven simulations…"*

The following sentence has also been added P7 L5-6 "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 (either 2050s or 2090s) and the reference period."

3. Section 2.2. The largest comment that I have on the paper is that there is insufficient information/discussion to convince the reader that the WaSim hydrology model does a good job of representing the catchment. Trust in this model is essential for the uncertainty analysis and conclusions of the paper. There is a brief mention of Nash Sutcliffe values at p4 L20, but the addition of a hydrograph plot showing modelled/observed values for some suitable period would make this more convincing. Especially given that Fig 3 shows significant under prediction of winter flow, and it is unclear what causes this problem.

Response: *A new figure showing the hydrographs has been added (i.e. Figure 3) to 2.2 with an additional table that lists some more performance statistics for several hydro gauges. The description of the model performance is now also covered in more detail P4 L31 – P5 L15: "While daily NSE values were lower for the three tributaries (Dart, Shotover and Matukituki River), monthly NSE values indicated a good performance (Table 2). For the Matukituki River the validation of both WaSiM versions revealed a substantially better performance (monthly NSE of 0.83 and 0.82, respectively) when compared to the TopNet based modelling study of Gawith et al. (2012) (monthly NSE of 0.68). For Chards Rd the performance of both 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 hydrographs of WaSiM-Anderson and*

WaSiM-Tindex (Figure 3) 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 1994) and an overestimation during autumn (e.g. April-May 1994). 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. Inaccuracies are generally larger with WaSiM-Tindex as can be seen in December 1994 (Figure 3). 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 performance of WaSiM-Anderson for the 1992-1996 period is substantiated by the daily NSE values which correspond to 0.91 and 0.87 for WaSiM-Anderson and WaSiM-Tindex, respectively."



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)

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.

Response: *See above and the following sentence (P13 L23-25) "Model parameter uncertainty was not accounted for in this study but should be part of future work, which could help to understand and potentially improve misrepresentations in the historic streamflow regime."*

P10 L20-31:" 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.

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.

4. Section 2.3. Worth noting that the climate scenarios used for New Zealand have now been superseded by 6 RCM*4 RCP Scenarios with CMIP-5 GCMs and a new bias correction that improves on quantile correction. See: http://ccii.org.nz/wpcontent/uploads/2016/10/RA1-Synthesis-report.pdf For the next paper perhaps!

Response: *We appreciate the suggestion of the author and have added the following comment to the conclusion section as part of future work P13 L12-15:" 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."*

5. Fig 6. I don't understand the comment about different y-axes.

Response: *Comment was unnecessary, agreed and deleted.*

6. Fig 9. Please include an explanation of what a radar chart shows.

Response: *Added: "Based on the permutations for each season the radar charts show the mean contributions [%] of the four model components to the overall uncertainty as well as the standard deviations [%]."*

7. p9 It would be useful to reference this paper: Hendrikx, J., Hreinsson, E.Ö., Clark, M.P. and Mullan, A.B., 2012. The potential impact of climate change on seasonal snow in New Zealand: part IâA˘Tan analysis using 12 ˘ GCMs. Theoretical and Applied Climatology, 110(4), pp.607-618.

Response: *Agreed paper has been added to document P10 L29-31 : "In a New Zealand wide study Hendrikx 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."*

General/additional response: *We have also made some minor modifications throughout the document in order to avoid confusion related to the terminology of runoff/streamflow. As modelled streamflow (m3/s) was converted to runoff (mm) for better comparison with other studies we have consistently replaced the term streamflow with runoff in the main document when we are referring to the actual results (as presented in the corresponding figures).*