Dear Editor,

please find enclosed a revised version of our manuscript now entitled "Snow water equivalent in the Alps as seen by gridded datasets, CMIP5 and CORDEX climate models".

We have addressed all points raised by the referees, performing supplementary analyses that in some cases were added in the manuscript or in the new Supplementary material, hopefully improving the overall quality and clarity of the work. In brief, the new parts cover:

- The clarification of the paper objectives, including the motivation on why we kept the section on future projections.
- The evaluation of the RCMs in the historical period, previously based on ERA-Interim driven runs, and now extended also to the GCM-driven models. A new figure (Figure S03) was added in the Supplementary material.
- Following the suggestions of Referee 2 we also analyzed the SNW distribution for different ranges of elevation, for all datasets (the references, GCMs, and RCMs). This analysis is now included in Fig. S04 of the Supplementary material.

We noted that for few comments of Revision 1, the page and lines indicated by the Reviewer do not match exactly the page/lines in the manuscript published on TCD. This was not a problem as we could easily associate the comments to the correct sentences in the text.

Overall we wish to thank the Reviewers for their constructive comments, which helped us to significantly improve the paper. We hope that this new version of the manuscript will be favourably considered for publication.

A point-by-point reply to the suggestions and comments of the Reviewers is reported below.

Reply to the comments of the Reviewer Yves Cornet

Research paradigms and hypothesis to be demonstrated

1. “The analysis of snow water equivalent in the Alpine region is thus a very challenging job because of spatial heterogeneity which is not taken into account in the 6 datasets used as reference (2 products derived from satellite observation and the 4 reanalysis), in the GCM and the RCM that have been compared. Nevertheless, regarding “real world” knowledge, figure 2 shows two maps provided by the HISTALP. I tell to the authors why they didn’t use this product as reference. I also tell them why they didn’t qualified HISTALP in a much more detailed way because I think it is a consistent representation of real world than the 6 ones selected as reference.”

The HISTALP gridded dataset provides a limited number of variables, including surface air temperature and total precipitation, and snowfall precipitation estimated from these. It is a good reference dataset but unfortunately it does not provide snow depth or snow water equivalent data, which are the focus of our study. The available variable have been stated in the text, at P8 L22-24. HISTALP temperature and precipitation climatologies have been shown in the paper and compared to the coarser resolution EOBS dataset to highlight the possible added value that high spatial resolution data can bring (the HISTALP spatial resolution is 0.083° lat/lon against 0.25° lat/lon of EOBS). Moreover, by comparing the two datasets, we highlight that uncertainties do exist *also* in *observational reference* datasets, and not only in climate models. Not unexpectedly, the uncertainty in temperature turns out to be lower than that found in precipitation. This is stated at P12 L1-4

2. "Moreover the inter-comparison of the Global and Regional climatic model and the reference datasets without knowing the "real world" evolution and its current situation is somehow disturbing for scientist. As a consequence the use of historical and predicted mean annual cycles from these models seems to me a very critical scientific paradigm which is non-pertinent. To conclude this section, I think that the comparison between models and the so-called references is probably interesting for climatic models developers. The analysis is thus acceptable with the exception of the section dedicated to the future evolution of the snowpack. It's thus absolutely inappropriate to present it as long as the demonstration of the reliability and the realistic spatial pattern of the SNW output of the models in Mountainous regions is not made. So, the question of major interest to be answered before this predictive operation with dangerous interpretative issues is the enhanced knowledge of the snowpack from finer observations by elaborating spatially representative sampling plan of the phenomenon and developing measurement methods enabling it to be implemented.

We agree that at the present state of knowledge, i.e. "without knowing the real world", projections of future snow depth are speculations. But the aim of this study is neither to state how snow water equivalent will evolve in the future nor to provide indications of the future state of snow resources. Instead we aim to show (i) how the uncertainty/spread found in the historical period project into the future, to assess the overall agreement on the relative changes with respect to each model climatology, and (ii) discuss whether the magnitude of the relative snow changes is similar in coarse and fine scale models. The spatial resolution is one order of magnitude finer in RCMs than in GCMs so that high elevation areas are resolved better in the former. We had two main questions in mind: (1) how does the resolution affect snow depth representation and its future changes in the Alpine environment? And, (2) is there any specific feature emerging in higher resolution projections, or are they indistinguishable from the lower resolution ones? This investigation is corroborated by a recent study comparing "bias corrected" and "non-bias corrected" snowfall projections of EURO-CORDEX RCM models (Frei et al. 2017). In that study bias corrected RCM snowfall was constrained to a snowfall reference dataset derived from 2 km resolution gridded temperature and precipitation data. According to that analysis, the relative change (RCP8.5 vs baseline) of the mean September-May snowfall is comparable whether applying or not the bias correction and the bias adjustment does not seem to have any significant effect on the trend.
We added in the introduction a sentence (P3 L34 - P4 L12) that states in a clearer way what the main purposes of this study are.

Methodological issues

Weighting procedure in the computation of RMSD, normalized variance and Pearson correlation.

- \- This procedure is described at p. 7. You assign a weight to each grid value given by the ratio between the area above 1000 m elevation and the area of the grid cell. You should give some arguments to justify that threshold and also to convince me that it is valid whole over the GAR.
  \- Further in the text (p. 12 l. 3, p. 13 legend caption) you explain the procedure in another way. I think that you should correct that to remain coherent through the whole paper to avoid ambiguity.

The weighting procedure mentioned by the reviewer in the first item above is applied only when snow water equivalent fields are spatially averaged over the Greater Alpine Region, i.e. in the plots shown in Figures 6 and 7, and not in the Taylor diagrams. The spatial averages of Figures 6 and 7 are intended to be representative of the mountains only, so we exclude the areas below 1000 m a.s.l. (we recognize that this threshold is arbitrary but we think that it could be appropriate in the GAR for focusing on high-altitude regions only) using this weighting procedure. The detail of the procedure is as follows: we weigh the snow water equivalent values at each grid cell by the area of the grid cell with mean elevation higher than 1000 m a.s.l. using a Digital Elevation Model at high spatial resolution (1 km), then the weighted values are spatially averaged over the domain of interest, the Greater Alpine Region. This is better explained in the manuscript at P10 L14-20.
For the Taylor diagrams, we calculated the root mean square error (RMSE), normalized standard deviation (NSD) and the correlation coefficient (R) over the full domain, without applying any weighting based on elevation. In this case the multiannual mean snow water equivalent was simply remapped onto the target grid conserving the snow mass from the original and the target grid cells. To this end we used a standard function incorporated in the CDO (Climate Data Operators) software mostly used in the climate community to handle climate model data in netCDF format. The CDO "remapcon" function performs an area-weighted remapping where the interpolation weights are based on the fractional area overlap of the original and target grid cells, following Jones, 1999. This methodology has been explained at P10 L7-14.

Jones, P.W. 1999, *Monthly Weather Review*, **127**, 2204-2210)

- - You should also provide a map in figure 1 () for instance with the spatial variation of this weight (area ratio).

Actually the "map of weights" is resolution-dependent so we should provide a map of weights for each dataset considered in this study. These maps would show for each coarse scale grid cell "the fraction of the area of the grid cell with mean elevation higher than 1000 m", where the topography is taken from the GLOBE Digital Elevation Model at 1 km resolution. This procedure has been better explained at P10 L14-20. We report below (Fig R1) two examples of maps of weights, the former referring to the CFSR reanalysis and the latter to the EC-Earth GCM. We prefer to not provide the maps of weights for each model in the manuscript because we think that this level of detail is too high for the general purpose of the paper.

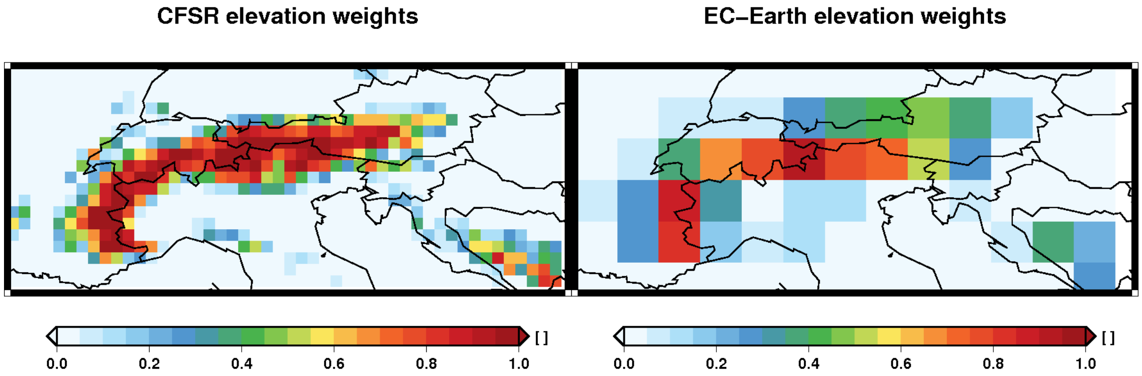


Fig. R1. "Map of weights" showing the fraction of each grid cell at elevation above 1000 m a.s.l. for the CFSR reanalysis and the global climate model EC-Earth. The reference topography is taken from the 1 km digital elevation model GLOBE (Hastings and Dunbar, 1999).

- - p. 7 l. 9-11 “This procedure allows for a fair comparison between datasets characterized by different spatial resolutions, without introducing uncertainties due to regridding”. I don’t totally agree with you. This procedure will enhance the importance of high area in the computation of “quality” parameters (Tailor diagrams). But high mountain zones are also very heterogeneous as low mountains zones. So the resolution difference effect will persist!!!

This comments reveals a misunderstanding. As explained before the weights are not used to calculate quality parameters (Taylor diagrams) for which we use the full domain. Instead, we used the weights approach to spatially average datasets characterized by very different spatial resolutions over the same domain (GAR above 1000 m a.s.l.), without interpolating the model data. The sentence has been rephrased in the manuscript at P10 L11-12.

> Interpolation. In your paper you use several words to describe the mathematical procedure used to change grid resolution (interpolation, reshaping, downscaling, remapping ...). When you reduce the ground sampling distance interpolation is the right terminology. When you degrade the resolution increasing g the ground sampling distance you perform a generalization of your geographical data and I think that this is a spatial aggregation method. You write at p. 10 : “To provide a fair comparison of the models and reduce the impact of the horizontal resolution on their performances, in particular on their spatial variance, each GCM is then compared to the MRM after having remapped each individual reference dataset onto the individual GCM grid, so that the reference is reshaped each time according to the model resolution. This approach allows for a fair comparison also for low resolution models.” Despite this statement I think that you should describe in a more detailed way the methods used to “reshape” your grids giving more information about “mass conservation” condition that should be verified.

We remapped the six reference datasets onto each climate model grid using a conservative remapping, in detail the CDO remapcon function (CDO 2015). This function performs an area-weighted remapping, where the interpolation weights are based on the fractional area overlap of the source and the destination grid cells, following Jones, 1999. Such interpolation weights applied to the source field allow to conserve the fluxes or water budgets from the source to the destination grid.
This procedure has been better explained adding some details at P10 L10-13.

CDO 2015: Climate Data Operators. Available at: http://www.mpimet.mpg.de/cdo

> Mean (central position statistics) computation.
> You compute the average of the references (satellite and Reanalysis products) .

- - The number of observations (6) is very small and one of these observations is obviously an outlier (20CR). Why did you this oulier?

  Thank you for this comment. We agree that after the assessment of its poor performance in representing SNW climatology, the 20CR reanalysis should not be considered as a “reference”. We repeated all the analysis excluding the 20CR from the Multi-Reference-Mean, and consequently figures 5,6 and 7 have been updated in the main text. The new procedure is explained at P13 L3-6;

  - This computation is performed on non- independent observation. For instance, it is quite clear for Global SWE Climatology and CFSR. So you give an exaggerated weight to those to datasets!!!

The interdependency of the Global SWE Climatology and the CFSR snow outputs has been better clarified in the text (P6 L6-7; P12 L11 and following). Both products integrate, but to different extents, the Special Sensor Microwave Imager (SSM/I) data. The Global SWE Climatology is specifically derived from Special Sensor Microwave Imager (SSM/I) data. The CFSR snow output is mainly based on the Noah land surface model first guess, and a daily snow analysis based on several inputs - among others Special Sensor Microwave Imager (SSM/I) data – is used to constrain the model first guess (Meng et al., 2012). In detail, CFSR snow depth/SWE are limited in the upper and lower boundaries by the snow analysis (it cannot be larger than twice and lower than half the snow analysis) but the temporal evolution of snow depth and SWE is determined by the Noah model. In conclusion, as the similarity between the two datasets is in the similar range of variability, we decided to include both.

> Same comments about the computation of the MMM (35 GCM or 7 HiRes GCM with ground sampling distance smaller than 1.25°).
>
> - GCM are probably not independent (see table 1) and some are probably highly correlated and have exaggerated weight!

We totally agree that the climate models are not independent from one another, and several previous studies (e.g. Knutti et al., 2013; Sanderson et al., 2015) focus on this issue. For example in Figure R2 below, now included in the Supplementary Material as Fig. S02, we report the spatial distribution of the DJFMA SNW in the 8 GCMs with horizontal resolution not coarser than 1.25°, referred to as "high-resolution" HiRes GCMs in the manuscript. These high-resolution models are CMCC-CM, EC-Earth, MRI-CGCM3, BCC-CSM1-1-M and four models from the CESM-family. Out of the four CESM-family models, one, CESM1-CAM5, shows a distinct behaviour. The other three (CESM1-BGC, CESM1-FASTCHEM and CCSM4) present very similar SNW patterns (Figure R2, last row), and similar RMSE, NSD and correlations values (Figure 5b in the main text). In order to have a model ensemble including models as independent as possible, we consider in the ensemble mean (MMM-HiRes) only one out of the three (CESM1-BGC). The analyses of figures 6 and 7 are then based on 6 high-resolution models, namely CMCC-CM, EC-Earth, MRI-CGCM3, BCC-CSM1-1-M, CESM1-CAM5 and CESM1-BGC.
In Figure 5a MMM-HiRes refers to the multi-model-mean of these 6 models. This choice has been explained in text at P13 L17-19 and later on at P19 L22-28.

Concerning GCMs with spatial resolution coarser than 1.25° it is difficult to evaluate their degree of interdependence from the Taylor diagrams. However, owing to their overall poor performances in the representation of SNW, and not being the focus of the paper, the aspect of their interdependency is not investigated further (explained in the main text at P19 L28-30).

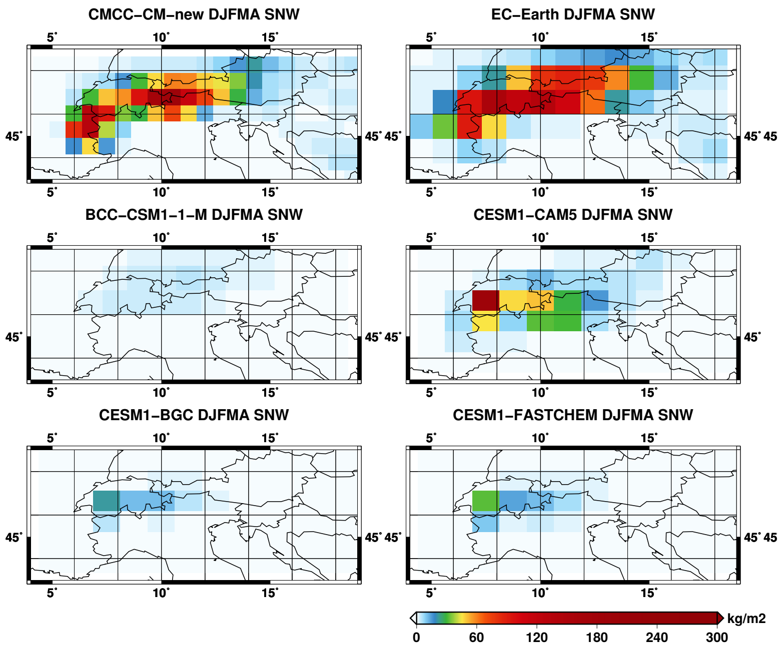


Fig R2: Multiannual mean (1980-2005) snow water equivalent in the GCMs with spatial resolution finer or equal to 1.25°.

What is the statistical pertinence of this aggregation method to determine central position statistics? Why a mean computation to "compensating extreme behaviours" (p. 10 l. 25) ? Why don't you compute the median (for instance) in this particular case (few and not independent observations, outliers).

We agree that it is interesting to explore the case in which the median is used as metrics, instead of the mean. In order to address this comment we explored two different approaches:

1) We considered as "reference" the median of the 5 datasets (NSIDC, CFSR, MERRA, ERAI/Land and 20CR), we calculated the median of CMIP5 models (full ensemble and HIRES ensemble) and we repeated the analysis of the Taylor diagram. The results are shown in Figure R3 . The median is shifted towards very small SNW values as NSIDC, CFSR and 20CR provide low snow. Consistently the normalized standard deviation of the MERRA reanalysis exceeds 2.5 and that of ERA-Interim lies outside the range of the plot, as well as for many climate models.
2) we consider as "reference" the average of 4 datasets (NSIDC, CFSR, MERRA, ERAI/Land), hence excluding the 20CR reanalysis from the "reference" statistics, as suggested by both reviewers. In this case we have a well-balanced ensemble, with NSIDC, CFSR showing low snow and MERRA, ERAI/Land showing large snow amounts. The results are shown in the new Figure 5a of the revised manuscript.

Considering the results of the two approaches we think that this second metric is the most appropriate to describe the "ensemble behavior" of the reference datasets and, therefore, it has been reported in the manuscript.

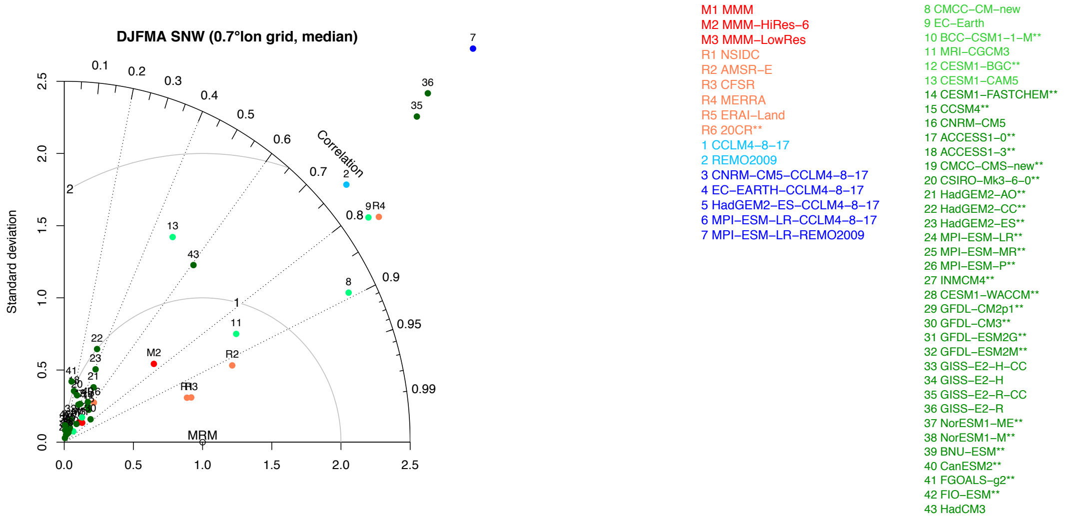


Fig. R3 : Taylor diagrams of the multiannual mean (1980-2005) of the DJFMA average snow water equivalent as described by climate models against the Multi-Reference-Median (MRM) calculated averaging NSIDC, CFSR, MERRA, ERAI/Land and 20CR climatologies. All datasets are projected onto the same reference grid at 0.7°lon.

Absence of spatial (geographical) analysis of the differences between the various spatial grids. To compare the grids you use 3 “quality” parameters reported in the Taylor diagrams. Even if Pearson-r is a measure of the association between variables and allows a global comparison of spatial patterns, I think that a spatial (or geographical) analysis of residuals (or the differences) is recommended to understand the effect of spatial localization. Doing so yous should be able to improve the discussion of some climatological factors that are not integrated in the same in the same way in the models and related to air mass circulations for instance: North-South of the Alps – humid and cold air-mass flow from the North or East-West - continentality and humid air-mass flux coming from the Adriatic towards South Eastern Alpine and Pre-alpine domain) on eventual systematic and variable biases.. This spatial analysis should thus be done for some specific and well-chosen models.

Here the referee states that “To compare the grids, the spatial (or geographical) analysis of residuals (or the differences) is recommended”, but actually this information is already contained in Figures 2,3,4 which show the differences between the various datasets (reanalysis, RCMs, GCMs) and the reference climatologies (EOBS for precipitation and temperature, NSIDC for snow water equivalent). Particularly important are, in our opinion, temperature and precipitation biases, that clearly show, for a given model, possible weaknesses related to the representation of the air mass circulation. These plots are commented in the corresponding section 4.1.1-4.1.3

Some specific comments

p. 1 l. 2. I’m very surprised about your conception of high spatial resolution in this abstract and in the whole text. For remote sensors (you use satellite data) hectometric and higher ground sampling distance corresponds to low and very low spatial resolution which don’t allow any description of bio-geo-physical processes on the Earth surface characterised by very high spatial frequency that are typical of mountainous area and especially the spatial variability of snow cover characteristics !

We agree that the definition of high resolution depends on the context. In the abstract P1 L2 (“high resolution, regional, observation-based gridded datasets”), “high resolution” refers to the typical spatial scales at which snow processes occur, i.e. less than 1 km. This has been clarified in the text

(P1 L2) Later on, when speaking about the resolutions of global climate models, the concepts of "high" and "low" resolutions refer to the typical horizontal grid size of the state-of-the-art numerical climate models (CMIP5), ranging from 70 to 400 km. In this case "high resolution GCMs" are those with grid size equal or finer than 1.25° (about 125 km). We added in the introduction (P3 L3-6) a sentence to clarify these definitions.

p. 1 l. 20 "The shift of the 0_C isotherm to higher elevations ..." Is it demonstrated overall on the GAR ?

Yes, because of an overall increase of surface temperatures (see i.e. Gobiet et al., 2014; Hantel et al., 2012; Serquet et al 2011; Beniston, 2003). We added the references in the text at P2 L7

p. 1 l. 22 "...decrease in the solid-to-total precipitation ratio in low- and mid-altitude mountain areas." What do you mean by low and mid- altitude? Does that definition depends on the climatological sub- domain within the GAR ?

At this point we are presenting a general picture, not focused on the Alps, and with "low and mid-altitude" actually we intend "areas with temperatures closer to the melting point". We have better specified this in the text at P2 L5-8 , thank you.

p.3 l. 14 What do you mean by large scale? The notion of scale in your document is somehow perturbing for cartographers and geographers that are specifically doing multiscale spatial analysis (see also p. 3 l. 14, p. 16 l. 1 for instance)! A map with a scale of 1:10000 is a large scale map that allows the representation and analysis of local physical phenomenon with small autocorrelation distance (high spatial variability). At the contrary a map with a scale of 1:1000000 is a small scale map that allows the representation and analysis of global phenomenon.

This sentence was present in a preliminary version of the paper but it has been removed in the version published on the online TCD http://www.the-cryosphere-discuss.net/tc-2016-280/tc-2016-280.pdf. Interestingly, "large-scale" and "small-scale" have opposite meaning in cartography (as pointed out by the reviewer) and in climate/geophysical fluid dynamics, where the large scales are those with the largest spatial extent and the small scales are those with smaller spatial extent. Curious discrepancy (in fact, opposite meaning) of terms in two neighbouring fields of research.

p. 6 l. 31-32 "Global climate models, also the most spatially resolved ones, do not take into proper account elevations above 1500 m a.s.l. over the GAR."
It's really a critical issue because it seems that "a very weak increasing trend towards heavier snowfalls has persisted since the 1960s" until 1999 in the Swiss Alps above the altitude of 1300 m as demonstrated by LATERNSER and SCHNEEBELI (2003, DOI: 10.1002/joc.912), for instance. But this research emphases the snow cover extent using low spatial resolution AVHRR images and you correctly state that satellite products provide a reliable picture of snow cover extent which is not the case for snow depth or snow water equivalent (p. 2 l. 20 and 21).

Yes, moreover, the period over which those trends are calculated (1931-1999) does not consider the last 17 years, generally characterized by low snow.

p. 10 l. 2 "... arbitrarily chosen ..."
This is not an acceptable. You should provide a scientific justification!

We chose the ERA-Interim Land grid as it has intermediate resolution between RCM and GCM grids. This explanation has been added in the methodology.

p. 10 l. 9 “... a wider distribution of SNW values, ...”→“... a wider statistical dispersion of SNW values, ...” (if I understand correctly)

It has been corrected in the text, thank you.

p.10 l. 4 “This second approach allows to minimize the impact of the horizontal resolution on the performances of GCMs.”
To Be Rewritten see next comment.

p. 10 l. 18 “... reduce the impact of the horizontal resolution on their performances ...”
I guess “their” refers to the models, then this sentence is not true. The impacts of the horizontal resolution on the models performance will not be reduced performing the reshaping of the reference datasets at the resolution of each GCM. To Be Rewritten.

Yes, we agree. We have changed the sentence (P19 L10-14) which now reads: “An alternative approach has been devised to provide a fair comparison of the GCMs. Each GCM is compared to the MRM after having conservatively remapped each reference dataset onto the individual GCM grid, so that the reference is reshaped each time according to the model resolution. This approach allows for a fair evaluation of the GCM at the model's grid, regardless of its resolution.”
Thank you for the comment.

p. 10 l. 28 “... of at least 1.25° ...”→“... finer than 1.25° ...”
In the document the concept of resolution is confused with that of Ground Sampling Distance!

Thank you. We have corrected it in the manuscript. In climate models, “resolution” refers to the physical distance (meters or degrees) between two consecutive gridpoints, in latitudinal, longitudinal or vertical direction, on the grid used to compute the equations” (IPCC, 2013)

p. 15 l. 3 “... wet precipitation bias ...” Pleonasm! -→“... overestimated precipitation ...” or “... positive precipitation bias ...”

Corrected, thank you.

p. 16 l. 23-24 “At global scale, the spread over mountain regions has been estimated to be several times larger than over midlatitude regions (Mudryk et al., 2015).”
I don’t understand why you compare midlatitude regions to mountain regions. The Alps are in a midlatitude region. You should complement the qualification of midlatitude regions!

Yes, we changed into “non-alpine midlatitude regions”, thank you

Comments on the document’s form (text, units, figures ...)

Units must be controlled:

- - p. 3 l. 16 “~80° km spatial resolution” ??
- - p. 8, l. 20 “105kg/m3 is not consistent with the unit used to describe the SNW in the reference

  datasets and the GCM (figure 2 for instance→kg/m2). It seems that you did this unit conversion to compute the mean annual cycle (figure 5 p. 13)

p. 9 figure caption 2 “... with horizontal resolution higher than 1.25°.”→“... with horizontal resolution finer than 1.25°”

p. 9 figure 2 “Panels (j,k) report the multiannual mean of the DJFMA accumulated snowfall derived from the HISTALP dataset. You should give a precision about the unit. I guess the unit in mm refers to the water equivalent volume per area unit! This value could be expressed using the same unit than the reference datasets and the GCM (kg/m2) assuming that 1 mm corresponds to 1 l/m2 ~ 1 kg/m2 !

p11. figure 3

- \- Labels are not readable even for points corresponding to cases with large NSD !!
- \- At that scale the large amount of points near the origin must be grouped in one class with a legend identifying all the point (dates or model) in the cluster.
- \- Colour is not the best graphical variable to identify the signification of the point reported in the legend and it is probably not necessary. If points are grouped combine in one class and write the composition of the class. If points are dispersed then then label is sufficient!!!

All the above technical comments have been accepted and modified in the manuscript accordingly. Thank you.

Reply to the comments by Reviewer#2

Reviewer: "In this paper, the authors assess the snow water content in the Alps as represented in several atmospheric reanalyses, ERA-Interim-driven (and to a lower extent CMIP5-driven) regional atmosphere models (EURO-CORDEX), and numerous CMIP5 models. I appreciate the large amount of datasets analysed in this study, however I have several concerns with the paper in its current form, and I think that a major overhaul is required before publication.

First of all, the aim of the present study is not clearly stated. Does the paper aim to provide projections of snow water equivalent for end users (ecologists, road managers, ski resort), or does it aim to assess the models fidelity in order to point out limitations in our ability to project future snow water equivalent or for any other purpose? The aim should be better explained, and this should also be used to choose and justify which diagnostics are shown in this paper (e.g. why evaluating ERAinterin-driven RCMs in section 4.1 and 4.2 and additionally evaluating CMIP5-driven RCMs only in section 4.3?)."

Reply: The clarification of the aims of this paper was also provided as a response to Reviewer#1 (question #2). Now the objectives of the paper are clearly stated in the introduction (P3 L26 – P4 L12). This paper does not intend to deliver snow water equivalent projections for end users: without a proper *absolute* validation of the accuracy of the model, future projections would be pure speculation. Instead this paper aims to show and point out the strengths and limitations in the current knowledge of snow water equivalent characteristics at regional scale.

In brief the main objectives are:

- to assess the uncertainties in the characterisation of current snow water equivalent in the GAR, from both satellite/reanalyses and climate models.
- to explore how the current model uncertainties project into the future.

For the first objective we need to evaluate ERA-Interim-driven RCMs and, ideally, the AMIP simulations of the CMIP5 experiment, as pointed out by the referee (thank you for the suggestion). Nonetheless, out of the 6 high resolution GCMs considered in this study, only two, CMCC-CM and MRI-CGCM3, have run AMIP simulations for the CMIP5 experiment (check in March 2017) and none of them is currently available for the download, apparently owing to issues with the servers. As of today, march 29th, we could not retrieve those data. At this stage it is impossible for us to evaluate the 2 AMIP runs.
For the second objective, we need GCM-driven RCMs and fully coupled GCMs. The scope of the manuscript is now better explained in the introduction.

"I also have a concern with the first diagnostics shown in this paper, i.e. the anomalies/biases represented on maps (Figs. 2-4). First, how are the datasets re-gridded prior to compute the difference? Furthermore, as the models (including reanalyses) miss the tail of the elevation distribution (as indicated in Fig.1), it is expected that they cannot account for high snowfalls observed in high-elevation areas. It seems to me that an alternative/complementary diagnostic would be to plot the snow water equivalent distribution per elevation bin. It would indicate whether the models behave well given their grid topography. I guess that the remapping used to build the Taylor diagram in Fig.5c partly addresses this, but it is not sufficient. In my opinion, this could replace sections 4.2 and 4.3 which I don't find very informative."

Throughout the paper the datasets are regridded using conservative remapping. This remapping allows the conservation of the quantity (SNW) from the original to the output grid.
Remapping methods do not change the original resolution of the datasets, so models and reanalyses

that do not represent the tail of the elevation distribution are not expected to represent high snowfalls observed in high-elevation areas. As suggested by the reviewer we produced a plot representing the snow water equivalent distribution per elevation bin (Fig R4). Reanalyses represent elevations up to 2000-2500 m; CMIP5 models generally represent elevations up to 2000 m; RCMs describe high elevation areas up to 3000 m. This plot clarifies what elevation ranges are represented in each dataset, thus it has been included in the Supplementary material in Figure S04. However this analysis does not show how close the modelled and the reference SNW patterns are, in terms of point-by-point correlation, mean error and variance. This information is instead given in the Taylor diagrams, which in our opinion provide much information in a concise way and, in our opinion, they cannot be replaced by the plot of SNW per elevation ranges alone.

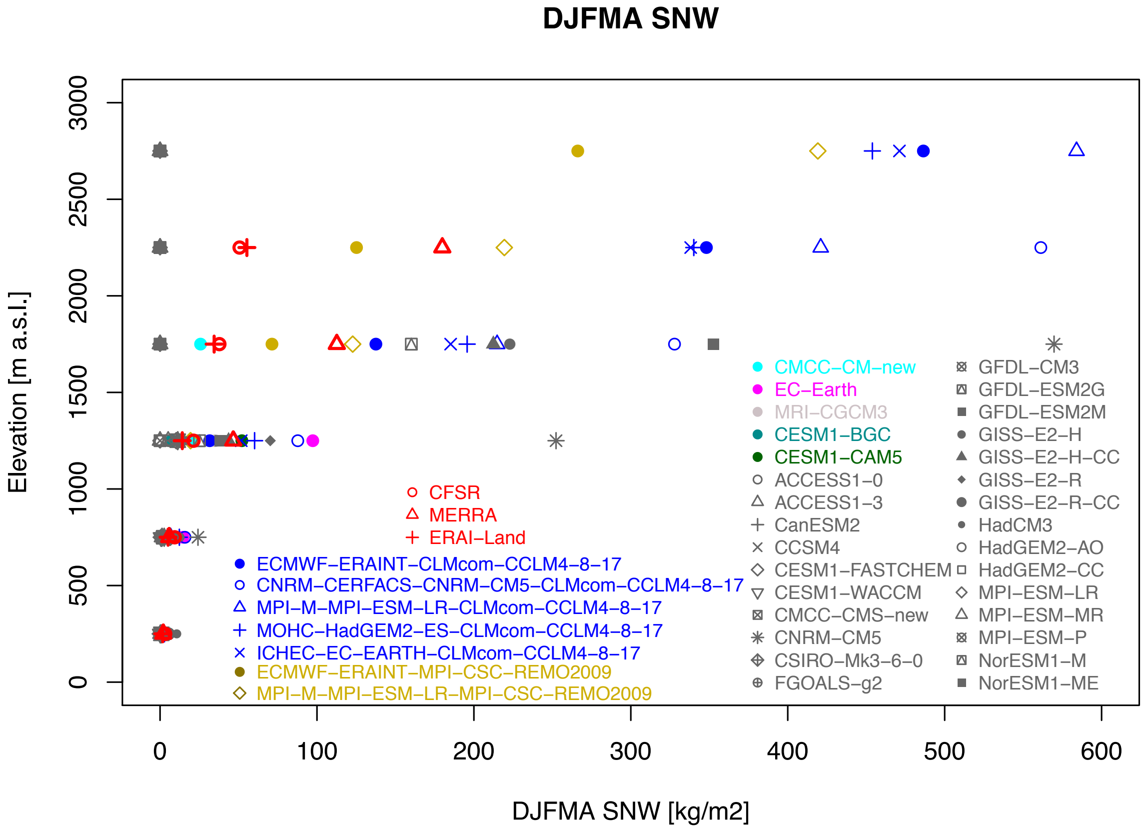


Fig R4. Multiannual mean DJFMA SNW in the Greater Alpine Region spatially averaged over different elevation ranges (500 m wide). The elevation is derived from the topography of each model (reanalysis).

“In addition, despite the limitations mentioned for snow water equivalent derived from passive microwave satellite observations, no attempt is made to discuss the validity of such products. The minimum would be to compare the two datasets described in section 2.1 over their common period. Of course, comparing to more datasets or to in-situ measurements would be even better. Is there any evidence that these satellite datasets are more reliable than the other datasets?”

The two satellite datasets AMSRE and Global SWE have been compared over their common period (2003-2007) and the plot has been integrated in Fig. 2f. We thank the reviewer for this suggestion. Given the relative short period of overlapping (5 years) we did not investigate further the time series, but we reported and discussed previous validation papers (Section 2.1 P5 L15-21)

“I have a possible concern with the choice of the simulations presented in the manuscript. ERAinterin-driven RCMs are more similar to AMIP models (atmospheric- only GCMs driven by observed SSTs) than to CMIP models, so in section 4.1 and 4.2, I think that comparing AMIP GCMS to ERAinterin-driven RCMs would make more sense. Then, in section 4.3 and 4.4 where the CMIP5-driven RCMs are evaluated, it makes more sense to compare to CMIP5 models.“

As stated before AMIP simulations are provided for only 2 HiRes GCMs and they are currently not available for download. However, we added Fig S03 in the supplementary material presenting the biases of GCM-driven RCM, to be compared to fully coupled CMIP5 models

“I have several other specific comments:

- Abstract, l.11: replace “latest” with “fifth” (in a couple of years, latest won’t be clear).”

Done, thank you.

- 2nd and 3rd paragraph of the Introduction: there are also concerns related to snow itself (road & airport safety, ski resorts, . . .).

Thank you. We have mentioned in the text the impacts on winter tourism and we added 2 citations (Beniston et al., 2011, Rixen et al., 2011). We preferred not to mention airport and road safety because it is more related to extreme events, i.e. to temporal scales not covered by our analysis.

- Intro, l.25-26: “at relatively high spatial resolution” -> subjective, indicate a typical range.

Done, thank you.

- It would be interesting to discuss the reliability of satellite datasets in the Introduction. Note that the GlobSnow dataset is derived from satellite measurements but uses ground-based weather station data in the SWE retrieval.

Yes, we added:
- a discussion on reliability of satellite datasets in (Section 2.1 P5 L15-21)
- the fact that GlobSnow is based also on “surface measurements” (P3 L18). Thank you

- Section 2.3: Sabin et al. (2013) use LMDz as an atmosphere-only model (i.e. not coupled to an ocean), I don’t know how relevant this is to the CMIP models.

Actually also Davini et al 2017 does. At present state of the art, ultra-high resolution simulations are AMIP only.

- Section 2.3, last paragraph, remove “at ISAC-CNR”.
- Tab. 1: there should probably be a line between satellite products and reanalyses.

Done, thank you.

- Given that LMD is mentioned, I’m surprised not to see the IPSL models in the long CMIP5 list, but anyway, there are clearly enough models in this paper.

IPSL models provide the snow depth variable but not snow water equivalent. Being the focus of this study on the latter variable, IPSL models do not appear in the paper.

- Sections 2.3 and 2.4: mention what kind of outputs are used (daily means or monthly means?).

We used monthly means. This detail has been added in the text (P7 L20 and P8 L2)

- Section 2.4: what is a “non-reliable snow accumulation trends”? (and what is a reliable trend?).

Pixels masked as “glaciers” do not reproduce the snowpack evolution (accumulation and melting) but they continuously accumulate snowfall in time (without melting). “Non-reliable trend” refers to this behaviour and it has been clarified in the text (P8 L4-8).

- Section 2.5, about “The ability of climate models to properly reproduce snow water equivalent depends both on the accuracy of their snow schemes and on the reliability of the atmospheric forcings”: it actually depends on many kinds on biases in the regional model (e.g. radiation scheme, boundary-layer scheme, etc, all being able to eventually impact snow).

Yes, we see your point. Actually with “reliability of the forcings” we already include all possible biases due to the land-surface and atmospheric schemes. We have rephrased the sentence “The ability of climate models to properly reproduce snow water equivalent depends on the accuracy of their surface snow schemes and on the reliability of the atmospheric fields forcing the snowpack processes.”

- Section 2.5: what is the interpolation method for HISTALP and EOBS?

EOBS is kept at its original resolution (0.25° lat-lon regular grid). HISTALP has been conservatively remapped to EOBS grid, as all the other datasets, for the comparison in Figure 2. This has been explained in the corresponding Section 4.1.1

- Section 3 could probably be merged with section 2 into a “datasets and methods” section.

Actually we prefer to keep them separate to make the text more readable.

- Fig.2: why showing the relative precipitation bias (in %) while the temperature and snow biases are shown as absolute errors?

Mainly to be consistent with a previous study by Kotlarsky et al., 2014, presenting the same maps for the same models over the full EURO-CORDEX domain. Here we present a focus on the Alpine region.

- Fig.2: the caption “snow water equivalent in the EOBS observational dataset and the NSIDC Global Monthly EASE-Grid Snow Water Equivalent Climatology respectively” is misleading, it would be clearer at a first read to write that EOBS relates to (a) and (b) while NSIDC relates to (c).

Done, thank you.

- Section 4.1.1, about “In order to facilitate the comparison we present the differences with respect to a given dataset: the NSIDC Global SNW Climatology for SNW, since it is available for a longer period (1980-2005) than the other satellite product AMSR-E (2003-2011)”. Ok, but it is a pity not to compare these two products aver the common period, especially given that you have claimed that “we expand the study by Mudryk et al. (2015) by including additional global SNW gridded datasets obtained from remote sensing” in the Introduction.

As previously mentioned, we have compared the two SNW satellite datasets over their common period (2003-2007) and the results are reported in Fig 2f. Given the relative short period of overlapping (5 years) we did not investigate further the time series, but we presented and discussed two papers on the validation of the two satellite products (Section 2.1 P5 L15-21). Thank you for the suggestion.

- Section 4.1.2: I would not say that REMO2009 is much better than the other RCMs, there is a substantial warm bias all over the domain (except maybe just over the mountain range) that could explain the relatively lower bias in SNW compared to other RCMs. Also, I would replace “CCLM4-8-17 and REMO2009 models which present no issues” with “CCLM4-8-17 and REMO2009 models which present weaker biases than other RCMs”.

We agree that the performance of REMO2009 are comparable to other RCMs (please note that the plot in Fig 4m has been updated after finding an error in the computation of the DJFMA mean). We have better explained in the text (Sections 2.4 and 4.1.3) the “issues” in ALADIN53, HIRAM5 and RACMO22E models: “these models show continuous snow accumulation and no melting in glacier-masked pixels. As this feature hampers the regridding of the model fields and the calculation of spatial averages over the GAR” we did not consider them for investigating the annual cycle and its projected changes at mountain range scale.

- Section 4.1.3: what period is used for the CMIP5 models, 1980-2005 or 1850-2005?

1980-2005, as clarified at P13 L16-17.

- Section 4.2 and its Taylor diagrams. I don’t find the spatial correlation very relevant here, because it mostly relies on correlations between the topographies. Similar comment for RMSE and NSD.

We agree that the correlation coefficient R mainly reflects the *model* topography but we do not this this is a limitation because each model has its own topography, at its own resolution. It would have been meaningless if all models were using the same topography. In our case the objective is to measure the similarity between climate model climatologies (provided at different resolution) and a reference pattern. In such case RMSE, NSD and R provide, in our opinion, a good measure of this similarity.

- Why removing the worst RCMs in section 4.2?

We did not remove the worst RCMs but the models presenting pixels characterized by continuous snow accumulation and no melting, possibly areas masked as glaciers. As this feature affects the water budget and it hampers the regridding of the model fields and the calculation of spatial averages over the GAR, we retained only two RCMs out of the five for further investigation. This has been explained in the text at P8 L4-8

- I am a bit lost, why using CMIP5-driven RCMs to analyse the seasonal cycle in section 4.3 and not to evaluate the mean spatial patterns in sections 4.1 and 4.2?

Yes, we added the evaluation of the GCM-driven RCMs in section 4.1.3, with one additional plot (Figure S3) in the Supplementary material.

- Section 4.3: I would not call 20CR a “reference dataset”, it is a coarse atmospheric GCM only constrained by surface pressure and SSTs, probably more comparable to a coarse AMIP model. . .

Thank you for this suggestion. As already mentioned in the Response to Reviewer 1, in the revised version of the manuscript the 20CR reanalysis is not considered as a “reference” any longer. In fact, we repeated all the analyses excluding the 20CR one from the Multi-Reference-Mean, and consequently figures 5, 6 and 7 have been updated in the main text. The new procedure is explained at lines P13L3-6.

Snow water equivalent in the Alps as seen by gridded datasets, CMIP5 and CORDEX climate models

Silvia Terzago[1], Jost von Hardenberg[1], Elisa Palazzi[1], and Antonello Provenzale[2]
[1]Institute of Atmospheric Sciences and Climate, National Research Council of Italy, Corso Fiume 4, Torino
[2]Institute of Geosciences and Earth Resources, National Research Council of Italy, Via Moruzzi 1, Pisa
Correspondence to: Silvia Terzago (s.terzago@isac.cnr.it)

Abstract. The estimate of the current and future conditions of snow resources in mountain areas ~~depends on the availability of reliable, high resolution~~would require reliable, kilometer-resolution, regional observation-based gridded datasets and ~~of~~ climate models capable of properly representing snow processes and snow-climate interactions. ~~Owing to the~~ At the moment, the development of such tools is hampered by the sparseness of station-based reference observations~~, in~~. In past decades mainly passive microwave remote sensing and reanalysis products have been used to infer information on the snow water equivalent distribution~~. However,~~, however, the investigation has usually been limited to flat terrains as the reliability of these products in mountain areas is poorly characterized.

This work considers the available snow water equivalent datasets from remote sensing and from reanalyses for the Greater Alpine Region (GAR), and explores their ability to provide a coherent view of the snow water equivalent distribution and climatology in this area. Further we analyze the simulations from the latest generation regional and global climate models (RCMs, GCMs), participating in the Coordinated Regional Climate Downscaling Experiment over the European domain (EURO-CORDEX) and in the ~~latest~~ Fifth Coupled Model Intercomparison Project (CMIP5) respectively. We evaluate their reliability in reproducing ~~snow water equivalent~~ the main drivers of snow processes - near surface air temperature and precipitation - against the observational dataset E-OBS, and the snow water equivalent climatology against the remote sensing and reanalysis datasets previously considered. We critically discuss the model limitations in the historical period and we explore their potential in providing reliable future projections.

The results of the analysis show that the distribution of snow water equivalent and the amplitude of its annual cycle are reproduced quite differently by the different remote sensing and renalysis datasets, which in fact exhibit a large spread around the ensemble mean. We find that GCMs at spatial resolutions finer than 1.25° longitude are in closer agreement with the ensemble mean of satellite and reanalysis products in terms of ~~RMSE~~ root mean square error and standard deviation than lower resolution GCMs. The set of regional climate models from the EURO-CORDEX ensemble provides estimates of snow water equivalent at 0.11° resolution that are locally much larger than those indicated by the gridded datasets~~but~~, and only in few cases these differences are smoothed out when snow water equivalent is spatially averaged over the entire Alpine domain. ERA-Interim driven RCM simulations show a snow annual cycle comparable in amplitude to those provided by the reference datasets, while GCM-driven RCMs present a large positive bias. ~~The~~ GCMs and higher-resolution RCM simulations are used to provide an estimate of the snow reduction expected by mid-21st century ~~in the~~ (RCP 8.5 scenario~~is weaker in higher-resolution~~

~~RCM simulationsthan in GCM runs~~) compared to the historical climatology, with the main purpose of highlighting the limits of our current knowledge and the needs for developing more reliable snow simulations.

1 Introduction

The increase in surface temperatures (IPCC, 2013) has relevant consequences on high elevation regions, where snow is a dominant climatic feature (Diffenbaugh et al., 2013; Barnett et al., 2005). The shift of the 0°C isotherm to higher elevations results in a decrease in the solid-to-total precipitation ratio in ~~low- and mid-altitude mountain areas~~mid- and low-altitude mountain areas, where temperatures are currently close to the melting point (Hantel et al., 2012; Serquet et al., 2011; Beniston, 2003). In addition, higher temperatures may result in earlier snow melt and shortening of the snow cover duration. Finally, snow cover and its local-scale variability affect climate at larger scales through the snow-albedo feedback (Scherrer et al., 2012).

Changes in mountain snowpack are expected to have implications on water availability, in particular on the timing of the seasonal runoff, likely characterized in the future by earlier spring or even winter discharge and reduced flows in summer and autumn (Beniston and Stoffel, 2014; Diffenbaugh et al., 2013; Barnett et al., 2005), and on the timing of the groundwater recharge. Similarly, changes in the seasonality and amount of winter snow cover and spring snow melt can have significant impacts on mountain economies for winter tourism (Beniston et al., 2011; Rixen et al., 2011) and on mountain ecosystems, including high-altitude vegetation (Körner, 2003) and the population dynamics of animal species that depend on snow resources (Imperio et al., 2013).

For these reasons, reliable regional estimates of current and future expected changes in snow cover are essential to develop adaptation and management strategies. Detailed studies on the recent and projected impacts of global warming in snow-dominated regions are necessary to inform future management of water resources (Beniston and Stoffel, 2014; Stewart, 2009; Barnett et al., 2005) and to preserve essential ecosystem services for millions of people living in downstream areas. For such applications, the uncertainties associated with the future snow projections must be carefully estimated and the reliability of the model results should be assessed.

In order to evaluate ~~the~~ state-of-the-art Global and Regional Climate Models (GCMs, RCMs) and their future projections, as well as to improve the representation of snow processes in such models, reliable datasets, possibly at high spatial resolution and representing the local climate characteristics in orographically complex areas, are required. However, the density of surface stations measuring snow is currently insufficient to develop a global, reliable gridded snow water equivalent dataset based on in-situ measurements, ~~which calls~~ thus calling for the use of alternative sources of information on snow depth and mass, derived from remote sensing observations and reanalyses (Mudryk et al., 2015).

Satellite measurements have been shown to provide a reliable picture of the global snow cover extent at ~~relatively high~~ few hundred meters spatial resolution (Brown et al., 2010; Hall and Riggs, 2007) while the estimation of snow depth and snow water equivalent from satellite is typically performed at spatial scales of 25 km and it is more challenging (Salzmann et al., 2014, see also Sect. 2). Global reanalyses provide snow water equivalent fields at horizontal resolutions that are comparable (∼30 km in the zonal direction) or coarser than satellite products. Some reanalyses, such as ERA-Interim (Dee et al., 2011) and

NCEP-CFSR (Saha et al., 2010), assimilate surface snow depth measurements and satellite snow cover extent while others, such as MERRA (Rienecker et al., 2011) and 20CR (Compo et al., 2011), are not constrained by measurements and thus rely on the capability of their land-surface model component to estimate snow fields. Overall, one must be aware of the very different meaning of "high resolution" in remote sensing studies, where spatial resolution can be of a few meters, and in climate modelling and /or gridded datasets, where the highest spatial resolutions that can be usually achieved are of the order of a few kilometers.

To date, few studies have investigated the accuracy of satellite-based and reanalysis snow water equivalent (SNW) datasets against available observations, and very little is known on their performances in mountain areas. Clifford (2010), for example, compared the long-term global snow water equivalent climatology provided by the National Snow and Ice Data Center (NSIDC, Armstrong et al., 2005), derived from passive microwave instruments, to the ERA40 reanalysis (Uppala et al., 2005) and to the output of the global climate model HadCM3 (Gordon et al., 2000; Collins et al., 2001). The largest differences between the three datasets were found for the Himalayas and for the west coast of North America, likely owing to heterogeneity of the sub-grid topography. Globally, the GCM and the reanalysis were found to be in ~~higher~~ better agreement with each other than with the satellite product. The GCM and reanalysis fields displayed a similar climatological annual cycle in the northern ~~Hemisphere~~hemisphere, a thick snow depth over Eurasia and a thin one over Siberia, while the satellite data indicated a thin snow pack in Eurasia and a thick one in Siberia overstimating snow depth with respect to the available ground observations. Another recent study by Mudryk et al. (2015) widened the analysis of Clifford (2010) by investigating additional SNW global datasets derived from satellite and surface measurements (GlobSnow, Takala et al., 2011), from reanalyses (ERA-Interim/Land and MERRA), and from land-surface models driven by meteorological forcing. The spread among these products was found to be lowest and their temporal correlation highest in mid-latitude boreal regions, likely owing to the fact that snow cover is generally ubiquitous during the cold season and the atmospheric circulation (midlatitude winter cyclones) is well reproduced in the models. The largest spread was found in Arctic and alpine regions, where reanalyses are poorly constrained by surface observations and the uncertainty in the meteorological forcing is higher. Alpine regions ~~present~~ are characterized by an additional complexity due to steep elevation gradients and sub-grid surface heterogeneities that are difficult to represent in land surface models.

The present work is devoted to review the available snow datasets, and to quantitatively assess the uncertainties in the estimation of the snow water equivalent in ~~a~~ alpine environment. First, we expand the study by Mudryk et al. (2015) by including additional global SNW gridded datasets obtained from remote sensing and reanalyses, and we explore how these datasets represent the snow climatology over the Greater Alpine Region (GAR). Based on this analysis, we critically discuss the performances of state-of-the-art SNW products in an orographically complex area and we provide an estimate of the inter-dataset spread in the Alps. These results are used as a reference for evaluating the state-of-the-art climate models participating in the two major coordinated global and regional climate modeling experiments~~,~~: the 5^{th} Coupled Model Intercomparison Project (CMIP5, Taylor et al., 2012), providing global simulations at spatial resolution on the order of 100 km, and the Coordinated Regional Climate Downscaling Experiment over the European domain ~~(EURO-CORDEX, Kotlarski et al., 2014)~~(EURO-CORDEX, Jacob et al., 2014), providing regional simulations up to 12 km spatial resolution. For each model, we assess its ~~snow water equivalent climatology~~

~~against the satellite and reanalysis datasets, and we measure the agreement among the different ensembles. The discrepancies among reanalysis and climate model simulations are discussed in relation to possible biases in the main driver~~ ability to represent (i) the main drivers of snow processes~~-~~, i.e. surface air temperature and precipitation~~- with respect~~, compared to the observational dataset E-OBS~~. Finally we provide an overview on the projected changes of the Alpine snow water equivalent~~, and (ii) its snow water equivalent climatology compared to the ensemble mean of the satellite and reanalysis datasets.

At the present state of affairs, i.e. without a sufficient knowledge of real surface snow conditions, it is not possible to make any statement on the reliability of future snow water equivalent projections at mountain range scale. In this study, without pretending to assess how snow resources will evolve in the future, we show how the model uncertainty and spread found in the historical period project into the future to (i) assess the overall agreement on the relative snow changes (i.e. changes relative to each model climatology) and (ii) discuss the differences in the amplitude of the relative snow changes projected by mid-21st century under a high-range emission scenario (RCP 8.5) ~~, and we highlight the differences between the~~ by coarse (CMIP5~~coarse-resolution projections and the finer-resolution ones from CORDEX~~) and fine scale (EURO-CORDEX) models.

The paper is organized as follows: section 2 introduces the datasets used for the analysis, section 3 describes the area of study, discusses the representation of orography in the current generation regional and global climate models, and summarizes the methodology employed for the data processing; section 4 reports the results in terms of (i) snowpack distribution in remote sensing products, reanalyses and climate model simulations over the Greater Alpine Region during the last decades, (ii) inter-dataset spread in the representation of the annual cycle of snow water equivalent, and (iii) inter-dataset spread in the representation of the snow changes expected by mid-21st century in the RCP8.5 scenario. Sections 5 and 6 provide a general discussion of the results in relation to other studies and conclude the paper.

2 Datasets

2.1 Remote sensing products

Satellite sensors can provide a reliable picture of the snow cover extent while the estimation of the snow water equivalent is more challenging. Passive microwave methods are based on the difference in brightness temperatures in two microwave channels, typically corresponding to frequencies of 18 GHz and 36 GHz. These methods are unable to detect very thin snow layers (i.e. less thick than 15 mm, Hancock et al., 2013) and suffer from saturation above $\sim$250 mm SNW (Clifford, 2010). Snow estimates from satellite are also affected by metamorphism of snow grains and snow melt: large, plate-like crystals increase the scattering of radiation from the surface, and a shallow but dense snowpack can be misinterpreted as a thick one. Owing to its high emissivity, liquid water, either within the snowpack or at the air-snow interface, overwhelms the scattering by the snow cover and can cause underestimation of the snow thickness. Additionally, melt-refreeze processes during the melt season can cause spurious snow peak values (Hancock et al., 2013). The horizontal resolution of satellite brightness temperature measurements makes the snow estimates extremely challenging in complex terrain owing to the heterogeneity of snow properties at subgrid scale. An eloquent example is the European Space Agency GlobSnow product in which the alpine regions are masked out because of intrinsic poorer performances and limited possibility to validate the snow estimates with surface observations (Takala et al., 2011).

Notwithstanding these limitations, satellite products are commonly used to evaluate SWE as they offer a global view on snowpack characteristics for several decades. In the present study we consider the following satellite products available for our study area:

- Global Monthly EASE-Grid Snow Water Equivalent Climatology (Armstrong et al., 2005) provided by the National Snow and Ice Data Center (NSIDC): This dataset includes global, monthly satellite-derived snow water equivalent data from November 1978 through May 2007 at 25 km resolution (Equal-Area Scalable Earth Grid, EASE-Grid). The snow water equivalent is derived from a Scanning Multichannel Microwave Radiometer (SMMR) and selected Special Sensor Microwave/Imagers (SSM/I).
- AMSR-E/Aqua Monthly L3 Global Snow Water Equivalent (level-3) monthly data (Tedesco et al., 2004) from the the Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E) instrument on the NASA Earth Observing System (EOS) Aqua satellite. This dataset contains SNW data and quality assurance flags mapped to 25 km EASE-Grids from 2002 to 2011.

NSIDC global monthly SWE data has been evaluated over Russia using snow observations from the period 1979-2000 for March only, showing an average 12 mm bias, which means a bias of 10% or less if the mean SWE is 120 mm or higher (Gan et al., 2014). The evaluation of the AMSRE SNW daily product in complex topography (Mackenzie River Basin, Canada) against in-situ snow depth observations showed similar results, a mean absolute error ranging from 12 mm in the early winter season to 50 mm in the late winter season (Tong and Velicogna, 2010). The differencese among the two satellite products over the Alpine region in terms average snow water equivalent during the overlapping period have been analyzed and discussed in Section 4.1.1, Figure 2f.

2.2 Reanalyses

A clear advantage of reanalysis products over observation-based data is that they provide global, physically-consistent estimates of all atmospheric and land-surface fields of interest, mostly constrained by observations. The reliability of reanalyses is related to the density of the assimilated observations, thus it depends on the location, the time period and the variable considered. Reanalysis products, for example, are known to be poorly constrained by surface measurements in mountain areas where their uncertainty is larger than in other regions. Precipitation is treated differently in different reanalyses: in some cases it is a prognostic variable, i.e. it is generated by the atmospheric general circulation model and it is not constrained by observations (i.e. MERRA reanalysis, Rienecker et al., 2011); in other cases it is a prescribed forcing derived from global precipitation datasets (as in the case of CFSR and ERA-Interim/Land reanalyses). The reanalysis products considered in the present study are:

- Climate Forecast System Reanalysis (CFSR, Saha et al., 2010) by the National Center for Environmental Prediction (NCEP), a global, high resolution, coupled atmosphere-ocean-land surface-sea ice system reanalysis, covering the period

1979-2009 and providing, among other variables, SNW fields at horizontal resolution $0.3125°$ ($\sim$ 38 km at the Equator). CFSR uses two sets of observed global precipitation analyses as precipitation forcing, namely CMAP (a 5-day mean precipitation dataset at 2.5 degree latitude-longitude grid) and CPC (daily gauge analysis at 0.5 degree lat-lon over land). CFSR snow fields are simulated by the land surface model Noah and constrained by the CFSR snow analysis. The snow analysis is based on the SNODEP model (Kopp and Kiess, 1996), which integrates surface observations, SSM/I-based detection algorithms and the NESDIS IMS North Hemisphere snow cover, based on in-situ and satellite data (Meng et al., 2012; Saha et al., 2010). Snow analyses are used to limit the upper and lower boundaries of Noah fields, that cannot be larger than twice and lower than half of the value provided by the analysis.

- Modern Era-Retrospective analysis for Research and Applications (MERRA, Rienecker et al., 2011) by the National Aeronautics and Space Administration (NASA), a global atmospheric reanalysis generated through the Goddard Earth Observing System Model (GEOS-5) atmospheric general circulation model and an atmospheric data assimilation system. MERRA covers the time period from 1979 through the present and it uses a grid of 1/2° latitude and 2/3° longitude with 72 vertical levels. Its land-surface model, Catchment (Koster et al., 2000), includes an intermediate complexity snow scheme with up to three snow layers describing snow accumulation, melting, refreezing and compaction in response to meteorological forcings (Stieglitz et al., 2001).

- ERA-Interim/Land reanalysis by the European Centre for Medium-Range Weather Forecasts (ECMWF), a global reanalysis of land-surface parameters at $\sim$80 km spatial resolution covering the period 1979-2010 (Balsamo et al., 2013). ERA-Interim/Land is the result of off-line simulations performed with the improved land-surface model HTESSEL (Balsamo et al., 2009) forced by the meteorological fields from ERA-Interim (Dee et al., 2011) and precipitation adjustments based on GPCP v2.1. ERA-Interim/Land re-scales ERA-Interim precipitation estimates on the Global Precipitation Climatology Project (GPCP) data to remove possible biases and add the constraint of observations on a monthly time scale (Balsamo et al., 2015). In fact, in the Alps ERA-Interim/Land has been found to reduce the dry bias present in ERA-Interim (see ~~Appendix ?? for details~~Fig. S01 of the supplementary material). At large scales, the correction on snowfall has been found to be small, owing to an overall good representation in the original ERA-Interim reanalysis (Brun et al., 2013). In ERA-Interim/Land snow density and snow depth are not constrained by data assimilation owing to limited availability of surface observations. In this way the accuracy of these variables relies purely on the capability of the HTESSEL land surface model to correctly reproduce the real fields. ERA-Interim/Land has been proven to provide good quality land snow mass analyses, owing mainly to the improvements in the single layer snow scheme, with enhanced parameterizations of snow density and revised formulations for the subgrid snow cover fraction and snow albedo (Balsamo et al., 2015; Dutra et al., 2010).

- 20th Century Reanalysis version 2 (20CRv2, Compo et al., 2011) by the NOAA Earth System Research Laboratory (ESRL) Physical Sciences Division and the University of Colorado CIRES Climate Diagnostics Center, providing a synoptic-observation-based estimate of global tropospheric variability spanning the time period from 1871 to 2008. It is derived using only surface pressure observations and prescribing monthly SST and sea-ice distributions as boundary conditions for the atmosphere (Compo et al., 2011). SNW fields are available at a spatial resolution of ∼ 1.875° (∼ 200 km in the zonal direction).

2.3 Global climate models

Global climate models (GCMs) are the main tools available to explore climate processes and feedbacks at global scales, and to make projections in future climate change scenarios. Owing to coarse-grid limitations, current GCMs resolve explicitly only the main snow processes while the snow physics at sub-grid scale is parameterized. In such conditions, the snow schemes used in GCMs are strongly simplified: they often treat snowpack as a single-layer over the ground surface and small-scale processes such as the refreezing of melted water within the snowpack and snow metamorphism are not properly taken into account (Steger et al., 2013).

Thanks to the availability of increasing computing resources it has been possible to run models at finer and finer spatial resolutions, thus permitting a more accurate representation of the topography in orographically complex areas (Davini et al., 2017; Sabin et al., 2013). Increased spatial resolution implies a more detailed view on the atmospheric forcings relevant for the mountain snowpack dynamics, i.e. altitudinal temperature gradients, precipitation distribution and phase, downward radiation, and the important physical processes could be better represented. As an example, the variable-resolution Laboratoire Meteorologie Dynamique (LMD) global climate model has been successfully employed to test the impact of the horizontal resolution on the representation of the monsoon over the South Asia (Sabin et al., 2013). They showed that the enhanced-resolution simulation at about 35 km greatly improves the representation of circulation features, the monsoon flow and the precipitation patterns with respect to the standard resolution model.

In the present study we consider the Global Climate Models included in the CMIP5 archive (http://www.cmip-pcmdi.llnl.gov/cmip5), as available in January 2015, providing the SNW variable at monthly resolution (Table 1) during both the historical period (1850–2005) and the projection period (2006–2100) under the Representative Concentration Pathways scenario RCP8.5 (Moss et al., 2010). We consider the ensemble member r1i1p1 for all models except for EC-Earth (Hazeleger et al., 2012) for which the SNW data were not stored in the CMIP5 archive and for which we used the ensemble member r8i1p1~~run at ISAC-CNR~~. The spatial resolution varies from model to model in a range from 0.75° to 3.75° longitude (∼80 to 400 km in the zonal direction, see Table 1).

2.4 Regional climate models

Dynamical downscaling of global climate models and reanalyses through regional models can potentially provide valuable information on the mountain cryosphere. Regional climate models are currently run at horizontal resolutions ranging from 50 km up to few km, allowing for a more refined representation of mountain topography and altitudinal gradients with respect to global models. Similarly to GCMs, RCMs snow schemes are strongly simplified with respect to dedicated snowpack models (Steger et al., 2013), so their main added value is to reproduce snow processes in high elevation areas, which are simply not represented in coarse grid GCMs.

In this work we consider all the RCMs participating in the EURO-CORDEX regional climate model experiment (Kotlarski et al., 2014) and providing the snow water equivalent variable at monthly resolution and at the finest available spatial resolution, i.e. 0.11° (Table 2). We evaluate the ERAInterim-driven runs, available for 5 models at the time we downloaded the dataset in October 2016, in order to assess the RCM bias when the RCM is driven by a realistic atmospheric forcing. Three models ~~present~~ show non-reliable ~~snow accumulation trends~~ trends (characterized by continuous snow accumulation and no melting) in a limited number of pixels - possibly areas masked as glaciers~~- so~~. As this feature hampers the regridding of the model fields and the calculation of spatial averages over the GAR we retained only two RCMs out of the five for further investigating the historical and the future simulations under the RCP 8.5 scenario (see Section 4.1.2 for details). Specifically one, the COSMO Climate version of Local Model (CCLM, Rockel et al., 2008) provides simulations driven by several different GCMs (namely EC-Earth, CNRM-CM5, HadGEM2-ES and MPI-ESM-LR) and thus it allows for investigating the uncertainty in the snow estimate coming from the large-scale driver. The other, REMO2009, provides simulations driven by the MPI-ESM-LR global climate model.

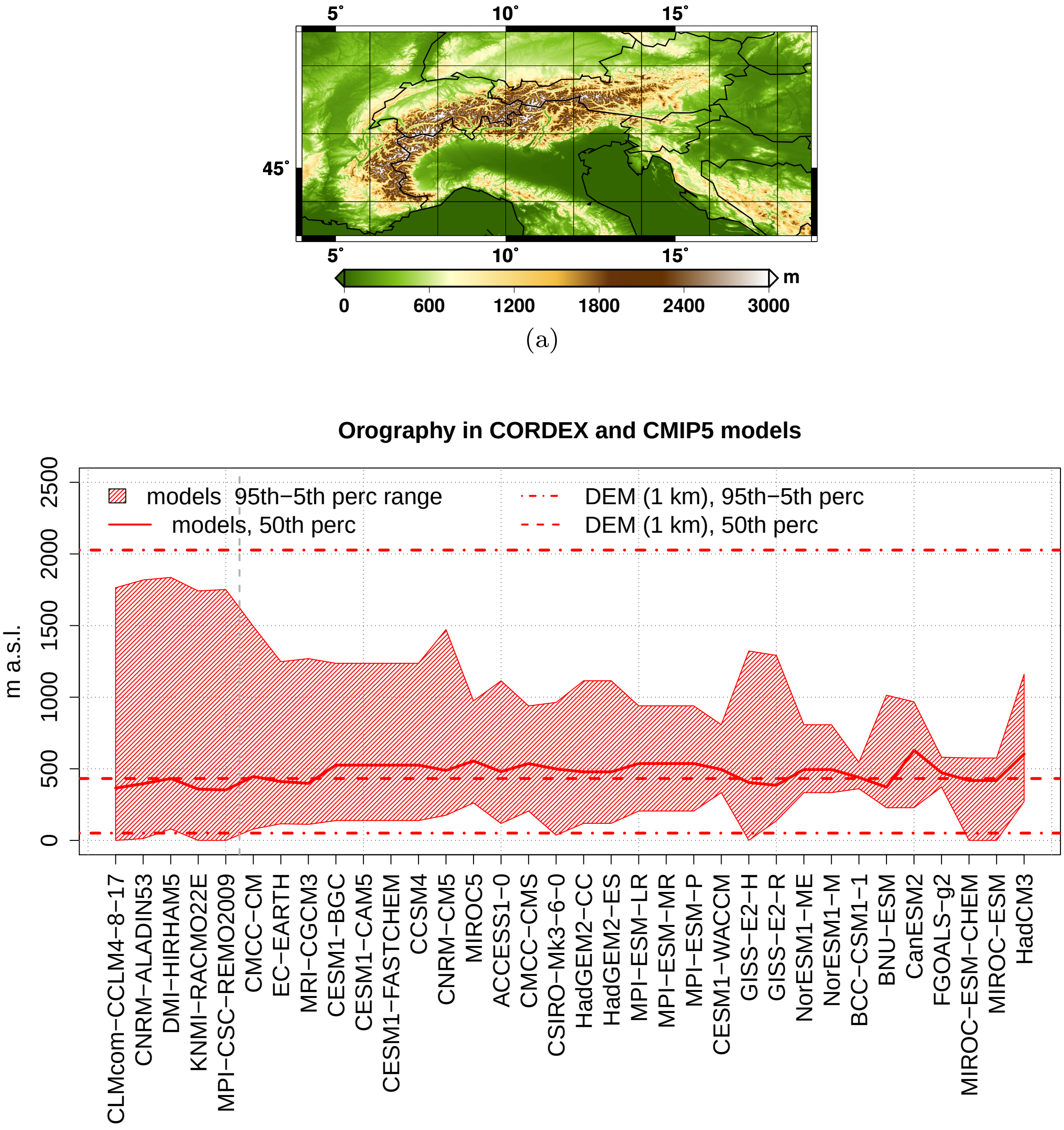
2.5 Observational datasets of air temperature and precipitation

The ability of climate models to properly reproduce snow water equivalent depends ~~both~~ on the accuracy of their surface snow schemes and on the reliability of the atmospheric ~~forcings~~fields forcing the snowpack processes. Near surface air temperature (TAS) and precipitation (PR) climatologies provided by the reanalyses and the climate models considered in this study are validated against two gridded observational datasets. Along the line of previous studies (Kotlarski et al., 2014) we consider the daily gridded ~~dataset EOBS~~ EOBS dataset (version 13, Haylock et al., 2008) at 0.25° resolution, based on the European Climate Assessment and Dataset station measurements.

In addition to this established and widely used reference, a second observational dataset specifically developed for the Alpine region, HISTALP ~~(Chimani et al., 2011), is also~~ (Auer et al., 2007; Chimani et al., 2011), is analyzed for comparison. HISTALP ~~allows to explore~~ provides monthly temperature and precipitation fields at 0.08° spatial resolution, and it based on surface measurements. Owing to its higher spatial resolution, HISTALP allows to explore such variables with a finer detail with respect to EOBS~~owing to its higher spatial resolution (0.08°)~~.

3 Domain and Methods

The study domain is the Greater Alpine Region (GAR, Auer et al., 2007), extending in the range 4–19° E, 43–49° N (Fig. 1a). The complex orography of the area and the heterogeneous pattern of steep slopes and valleys make a proper representation of the climate features challenging from both an observational and a modeling point of view. As an example, Fig. 1b points out how the topography is represented in the 1-km GLOBE digital elevation model (Hastings and Dunbar, 1999), in the CORDEX ERA-Interim driven regional climate models and in the CMIP5 global climate models, in terms of median and 95th percentile of the distribution of elevation. The median elevation is well reproduced by all models while the lowest and highest elevations are progressively cut out as the model spatial resolution decreases. While RCMs are closer to the expected values, global climate



(a)

(b)

Figure 1. (a) Orography of the Greater Alpine Region (4-19° E; 43-49° N), as in the GLOBE 1 km digital elevation model (DEM). (b) The 95^{th}, 50^{th} and 5^{th} percentiles of the elevation distribution in the DEM (dash-dot and dashed lines, respectively), compared to the corresponding values obtained from the CORDEX and CMIP5 model orographies. RCM and GCM models are ordered along the x-axis from finest to the coarsest spatial resolution. RCMs and GCMs are separated by a vertical dashed line.

models, also ~~the most spatially resolved ones~~those with the finest spatial resolution, do not take into proper account elevations above 1500 m a.s.l. ~~over~~ in the GAR. This limitation has to be considered when analyzing GCM outputs over mountain areas since the world reproduced by the global models has a smooth orography and simplified physical processes.

In this paper we explore the degree of agreement (i) among the reference datasets illustrated in Sect. 2.1 and 2.2, (ii) of the CORDEX and CMIP5 models compared to the ensemble mean of the reference datasets and (iii) among the different climate model ensembles, by ~~visually inspecting the DJFMA~~ inspecting the December to April (DJFMA) mean TAS, PR, and SNW climatologies.

The model performance with respect to the reference snow water equivalent datasets is quantified using Taylor diagrams, which provide a concise statistical summary of how well patterns match a given reference in terms of their correlation (R), root-mean-square difference (RMSE), and ratio of their variances (NSD) (Taylor, 2001). In order to compare point by point datasets built on different coordinate reference systems and with different spatial resolutions, all datasets are reprojected onto a common grid, the ERA-Interim/Land 0.7° longitude grid, chosen because of its intermediate resolution between global and regional climate models. Global climate models are also evaluated at their own resolution, comparing each model to remote sensing products and reanalyses upscaled at the climate model grid. This second approach allows to reduce the impact of the horizontal resolution on the performances of coarse scale climate models. Spatial interpolations are performed via conservative remapping (Jones, 1999), using the Climate Data Operators software (CDO, 2015).

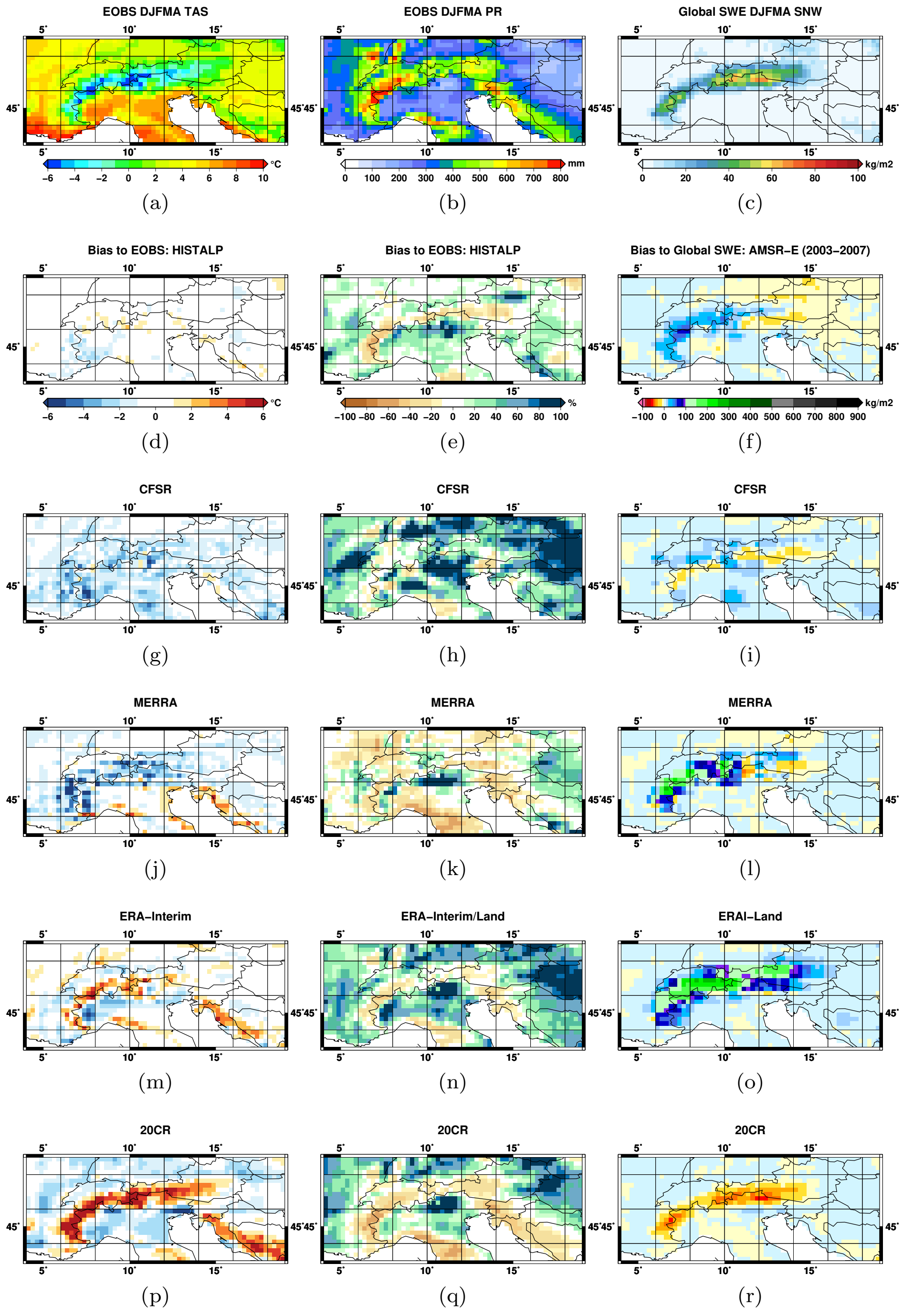
Assessments of the SNW characteristics at the scale of the mountain range (Figures 6 and 7) are obtained by spatially averaging the snow water equivalent over all areas above 1000 m a.s.l. in the GAR. To take into account the mismatch between the model topography and the real one, we use the datasets at their native resolution and weight the values by the fraction of each grid cell at elevation above 1000 m a.s.l as provided by the 1-km GLOBE (Hastings and Dunbar, 1999) digital elevation model; then the weighted values are spatially averaged over the domain of interest, the Greater Alpine Region. This procedure allows ~~for a fair comparison between~~ to compare datasets characterized by very different spatial resolutions ~~,~~ without introducing uncertainties due to regridding ~~(for futher details see Terzago et al., 2014)~~(see also Terzago et al., 2014, for further details).

4 Results

4.1 The spatial distribution of snow water equivalent in gridded datasets

4.1.1 SNW in satellite products and reanalyses

We first illustrate the spatial distribution of snow water equivalent in the satellite products and the reanalyses, hereafter referred to as the *reference* datasets, and we evaluate the differences among the reanalyses in relation to possible biases in the meteorological forcing. Figure 2 shows the multiannual mean (1980-2005) of SNW, near surface air temperature (TAS) and precipitation (PR) averaged (or accumulated in the case of PR) over the months from December to April. In order to facilitate the comparison we present the differences (or percent biases) with respect to a given dataset: the NSIDC Global SNW Climatology for SNW, since it is available for a longer period (1980-2005) than the other satellite product AMSR-E (2003-2011); EOBS observations for TAS and PR~~(for the latter we show percent bias)~~. All datasets are conservatively remapped on a 0.25° resolution regular grid. Biases are calculated over the period 1980-2005 except for AMSR-E, for which the period of overlap with the reference dataset is 2003-2007.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

(j) (k) (l)

(m) (n) (o)

(p) (q) (r)

Figure 2. Multiannual mean (1980-2005) of the DJFMA average (a) air temperature, (b) total precipitation from EOBS observational datasets and (c) snow water equivalent from NSIDC global SNW. Panels from (d) to (r) represent the bias of HISTALP, AMSR-E and reanalyses with respect to EOBS and NSIDC datasets, respectively.

Compared to EOBS, the alternative observational, high-resolution climatology from HISTALP (Fig. 2d-e) presents a similar temperature distribution, drier conditions at high elevations and wetter conditions at low elevations. This comparison is reported to highlight the fact that uncertainties are ~~greater~~ larger in precipitation than in temperature estimates, especially in mountain areas, and also observational datasets can exhibit biases with respect to each other.

Focusing on the snow water equivalent distribution, the NSIDC Global SNW climatology (Fig. 2c) shows maximum values of about 50 kg/m^2 over ~~Western~~ the western Alps and 70 kg/m^2 over Eastern Alps. If we consider the other satellite and reanalysis products we obtain a rather heterogeneous picture. AMSR-E (Fig. 2f), which is derived as well from remote sensing observations, presents higher values in ~~Wester~~ western Alps and lower values in the ~~Eastern~~ eastern Alps compared to the NSIDC climatology.

CFSR (Fig. 2g-i) shows TAS and PR patterns that are similar to EOBS over the Alpine ridge, and ~~SNW distribution~~ a SNW distribution that is similar to NSIDC Global SNW. The similarity in SNW ~~distribution~~ range of variability is probably due to the fact that ~~CFSR assimilates a snow mask derived from the same sensors as NSIDC, so these datasets are not~~ both products integrate, but to different extents, the Special Sensor Microwave Imager (SSM/I) data. The Global SWE Climatology is specifically derived from Special Sensor Microwave Imager (SSM/I) data. The CFSR snow output is mainly based on the Noah land surface model first guess, and a daily snow analysis based on several inputs, including, among others, the Special Sensor Microwave Imager (SSM/I) data - is used to constrain the model first guess (Meng et al., 2012). In detail, the CFSR snow depth/SWE is limited in the upper and lower boundaries by the snow analysis (it cannot be larger than twice and lower than half the snow analysis) but the temporal evolution of snow depth and SWE is determined by the Noah model. As a consequence, the two SNW datasets lie in similar ranges of variability, but except for this feature they can be considered as independent.

The MERRA Reanalysis (Fig. 2j-l) shows a thicker snowpack with respect to NSIDC Global SNW as well, especially over the Western Alps, as well as AMSR-E. The MERRA behavior can be explained by a cold bias over that area, partly compensated by drier conditions over the Alpine peaks.

ERA-Interim/Land (Fig. 2m-o) shows the largest SNW values, with peaks exceeding NSIDC values by more than ~~100~~ 200 kg/m^2. The SNW bias is not directly explainable in terms of biases in temperature and precipitation, which indeed go towards the opposite direction (~~slightly warmer and~~ warmer and slightly drier with respect to EOBS). This result suggests that ERA-Interim/Land high SNW values ~~are~~ can be ascribable to the snow scheme in use.

~~The MERRA Reanalysis (Fig. 2j-l) shows thicker snowpack with respect to NSIDC Global SNW as well, especially over the Western Alps, as well as AMSR-E. The MERRA behavior can be explained by a cold bias over that area, partly compensated by drier conditions over the Alpine peaks.~~

20CR (Fig. 2p-r) shows the lowest SNW values. Owing to its coarse spatial resolution, 20CR presents a warm and dry bias at high elevations and a cold and wet bias at low elevations, which in turn result in low snow accumulation and shallow snowpack over the mountain range. These simplified patterns can presumably be ascribed to an excessively smooth orography and highlight the limitations of the 20CR reanalysis in the representation of snow processes in mountain areas.

This analysis provides a quite heterogeneous picture on SNW and, despite the considerations on the biases of the drivers, it is not possible, at the present state of knowledge, to ultimately define which product is closest to the reality over the full GAR domain. ~~Therefore, for~~ For further analysis we ~~use the mean of all reference datasets(~~disregard the 20CR reanalysis owing to its poor performaces in this orographically complex region and the AMSR-E satellite products for its short period of availability. We consider as *reference* the mean of the other 4 datasets, i.e. NSIDC Global SWE, CFSR, MERRA and ERA-Interim/Land reanalyses. This Multi-Reference Mean ~~,~~ (hereinafter MRM) ~~calculated after interpolating~~ is calculated after conservatively remapping all the datasets to the 0.7° longitude ERA-Interim/Land grid.

~~Multiannual mean (1980-2005) of the DJFMA average (a) air temperature, (b) total precipitation and (c) snow water equivalent in the EOBS observational dataset and the NSIDC Global Monthly EASE-Grid Snow Water Equivalent Climatology respectively. Panels from (d) to (r) represent the bias of HISTALP, AMSR-E and reanalyses with respect to EOBS and NSIDC datasets, respectively.~~

4.1.2 SNW in ~~regional~~ global climate models

~~Figure 3 represents~~ Here we discuss in detail the DJFMA TAS, PR and SNW climatologies ~~as in ERA-Interim-driven RCM simulations averaged over the period 1990-2005. As in~~ provided by CMIP5 global climate models with spatial resolution equal or finer than 1.25° (Figure 3); coarser resolution GCMs are discussed further in Section 4.2.

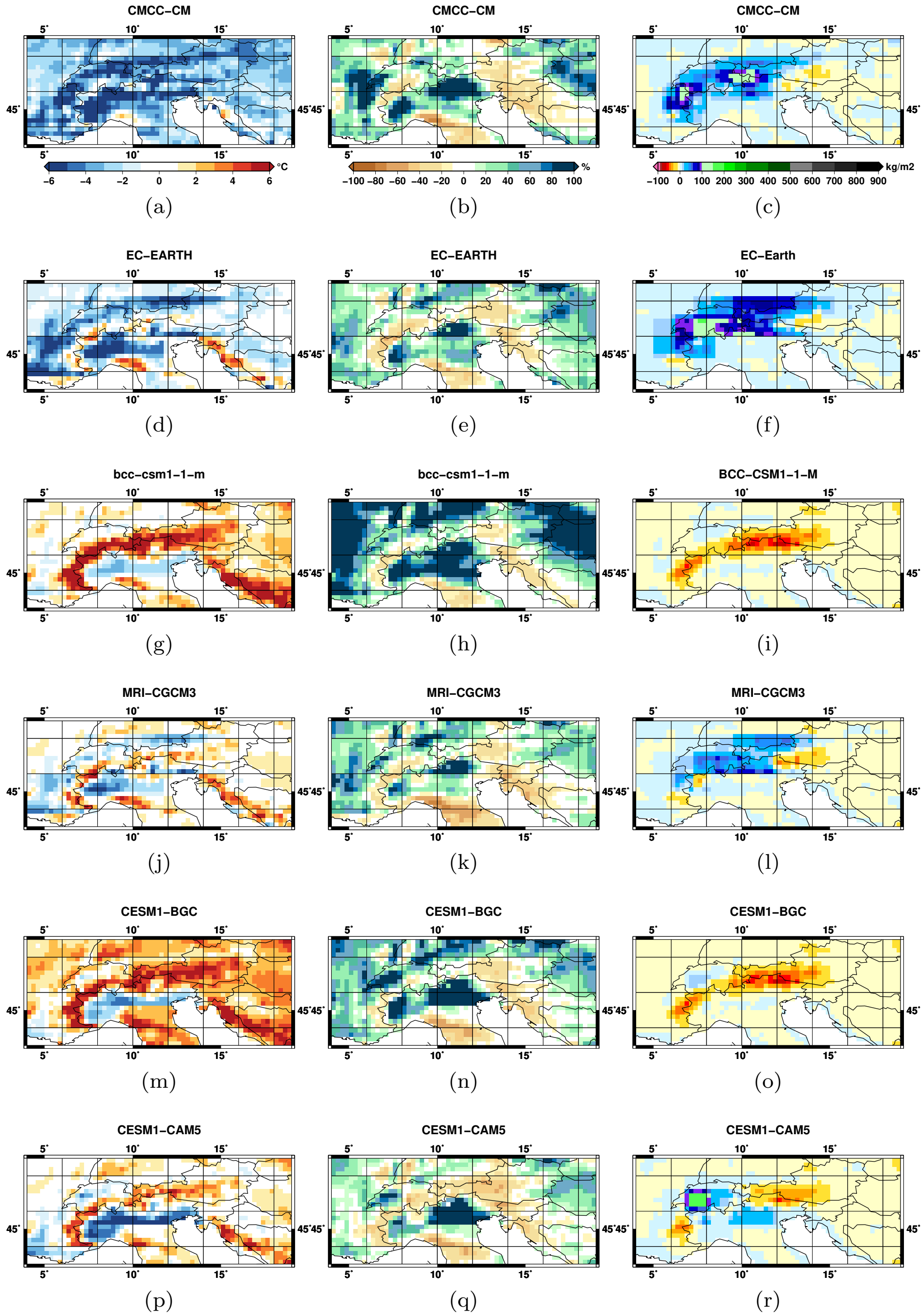
CMIP5 model biases with respect to EOBS and NSIDC references (Fig. 2~~we show the biases~~ a-c) are shown Fig. 3. The comparison period is 1980-2005. Among the 4 CESM-family models, namely CESM1-CAM5, CESM1-BGC, CESM1-FASTCHEM and CCSM4, three models present very similar climatologies so we consider here only one of them, CESM1-BGC, taken to be representative also of CESM1-FASTCHEM and CCSM4 (see Figure S02 and Section 4.2 for further details).

GCMs with spatial resolution finer than 1.25° show snow amounts which are comparable to those of the reference datasets over the Greater Alpine Region. Compared to the NSIDC SNW, CMCC-CM, EC-Earth and, to a smaller extent, MRI-CGCM3 and CESM1-CAM5 models, show thicker snowpack at the northern slope of the Alps and in Switzerland. A common feature of all datasets is a shallower snowpack over the eastern Alps, at the border between Italy and Austria. This spatial pattern, characterized by an East-West gradient, with shallower snowpack in the eastern Alps and thicker snowpack in the western Alps, is more resemblant to that provided by the AMSR-E satellite products rather than to the NSIDC Global SNW.

BCC-CSM1-1-M and CESM1-BGC show shallower snowpack than the NSIDC Global SNW, and higher temperatures with respect to the observational datasets. In these cases the warm bias in the model can explain a less abundant snowpack.

From this analysis the precipitation bias over the Alpine ridge seems comparable among the different high resolution GCMs. In fact, GCMs generally tend to a slight underestimation of winter precipitation at the ridges and to an overestimation at lower altitudes. This uniform behavior in the precipitation pattern suggests that temperature can be the leading factor which determines biases in the estimation of the snow depth.

4.1.3 SNW in regional climate models



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

(j) (k) (l)

(m) (n) (o)

(p) (q) (r)

Figure 3. DJFMA (first column) air temperature, (second column) total precipitation and (third column) snow water equivalent biases of the CMIP5 global climate models with spatial resolution equal or finer than 1.25° longitude with respect to the EOBS and NSIDC SNW climatologies reported in Fig. 2a,b,c.

Figure 4 shows the biases of ERA-Interim-driven regional climate models DJFMA TAS, PR and SNW climatologies with respect to the EOBS and NSIDC SNW ~~average fields~~references, all averaged over the common period 1990-2005.

All RCMs show SNW amounts several hundreds kg/m^2 larger than any other reference dataset (Fig. 2) at the mountain ridge and lower values at low elevations. Actually ~~some~~ extremely high values (shown in black) are non-reliable as they correspond to areas ~~where snow can accumulate indefinitely~~of continuous snow accumulation and no melting, possibly areas masked as glaciers in the models. Such gridpoints ~~present artificially high~~ show artificially high erroneous positive trends and they have to be discarded from the analysis. Despite these details, RCM snow estimates are much higher than those provided by the reference datasets, and these high values can be related to the fine representation of the orography ~~which~~ that allows, in principle, for ~~colder~~ lower temperatures in high mountain areas, not represented in coarse-scale reanalyses, for increased solid precipitation and longer snow pack duration.

In some cases the large SNW values in RCMs can be partly explained with cold biases (RACMO22E, ALADIN53) or wet biases (HIRHAM5) with respect to the observations. In other cases (~~REMO2009 and~~ CCLM4-8-17) ~~the~~ despite remarkable biases in some parts of the domain, the atmospheric forcings in correspondence of the mountain ridge are in ~~overall~~ better agreement with observations and they do not show relevant deviations from the reference climatologies, so the differences have to be attributed to the snow scheme in use and/or to the finer representation of the topography.

From the analysis of RCMs we can conclude that higher spatial resolution allows to better separate areas of snow accumulation and, consequently, to reproduce higher snow maxima in correspondence of mountain peaks.

For ~~further investigations we will mantain only~~ the CCLM4-8-17 and REMO2009 models~~which present~~ , which display no issues in the snow accumulation trends~~.~~

~~Biases in DJFMA air temperature, total precipitation and snow water equivalent in the CORDEX ERA-Interim-driven RCM simulations, with respect to EOBS and NSIDC Global Monthly EASE-Grid Snow Water Equivalent Climatologies reported in Fig. 2a,b,c.~~

4.1.4 ~~SNW in global climate models~~

~~GCMs with highest spatial resolution (finer than 1.25°, Fig. 4)present considerably lower SNW amounts with respect to RCMs and comparable amounts with respect to reference datasets.~~

~~Compared to NSIDC SNW, CMCC-CM~~, we investigated also the GCM-driven simulations (Table 2). GCM-driven CCLM4-8-17 climatologies possess a stronger negative temperature bias (CNRM-CM5, EC-Earth~~and, to a smaller extent, MRI-CGCM3 andCESM1-CAM5 models, show thicker snowpack at the Northern slope of Alps and in Switzerland. A common feature of all datasets is a shallower snowpack over the Eastern Alps, at the border between Italy and Austria. This spatial pattern, characterized by an East-West gradient, with shallower snowpack in Eastern Alps and thicker snowpack in Western Alps, is resemblant to that provided by the AMSR-E satellite products rather than by NSIDC Global SNW.~~

~~BCC-CSM1-1-M~~, ~~CESM1-BGC and CESM1-CAM5 show shallower snowpack than the NSIDC Global SNW, and warmer temperatures with respect to the observational datasets. In these cases the warm bias in the model can explain a less abundant snowpack.~~

~~Precipitation bias over the Alpine ridge seems comparable among the different high resolution GCMs. In fact, GCMs generally tend to a slight underestimation of winter precipitation at the ridges and to a slight overestimation at lower altitude. This uniform behavior in the precipitation pattern suggests that temperature can be the leading factor which determines biases in~~ HadGEM2-ES) and/or stronger positive precipitation biases (CNRM-CM5, MPI-ESM-LR) with respect to the ERA-Interim driven runs (Figure S03). These feature result in thicker snow water equivalent. In the ~~estimation of the snow depth~~case of MPI-ESM-LR-driven REMO2009 the temperature bias is comparable while the precipitation bias is larger than for the ERA-Interim driven runs. In conclusion, GCM-driven RCM simulations tend to suffer the biases already present in the driver GCM and to reflect them in SNW fields.

4.2 Global view on SNW products

In this section we provide a comprehensive view on all the previously considered SNW gridded datasets. The similarity of the SNW climatologies ~~shown in Figs. 2, 3 and 4~~ is quantified using the metrics of Taylor diagrams (Taylor, 2001). Figure 5a compares the spatial distribution of the DJFMA snow water equivalent, averaged over the period 1980–2005, for the Multi-Reference-Mean (MRM~~)~~, mean of ~~all reference datasets~~ the 4 reference datasets CFRS, MERRA, ERA-Interim/Land and NSIDC Global SWE) to which all other datasets are compared; the Multi-Model-Mean (MMM), mean of all ~~35~~ 36 CMIP5 models; the Multi-Model-Mean of the CMIP5 models with spatial resolution finer than 1.25° (MMM-HiRes, as in Terzago et al., 2014); the individual reference datasets; and the individual regional and global climate models. ~~In order to~~

First we compare datasets built on different coordinate reference systems and with different spatial resolutions ~~, two different approaches have been followed. First,~~ by reprojecting all remote sensing products, reanalyses and climate model outputs ~~are reprojected~~ onto a common grid, ~~arbitrarily chosen as~~ specifically the ERA-Interim/Land 0.7° longitude grid. ~~Alternatively, climate models are evaluated at their own resolution, comparing each model to remote sensing products and reanalyses upscaled at the climate model grid. This second approach allows to minimize the impact of the horizontal resolution on the performances of coarse scale climate models, and in fact it is applied to GCMs only.~~

~~Taylor diagrams of the multiannual mean (1980-2005) of the DJFMA average snow water equivalent as described by climate models against the Multi-Reference-Mean (MRM): (a) all datasets are projected onto the same reference grid at 0.7°lon; (b) the climate models are kept at their original resolution and the reference datasets are remapped onto the grid of each model.~~

~~Figure 5 a shows the results for the first approach. It~~Figure 5 provides an evaluation of the individual datasets with respect to the Multi-Reference-Mean, all resampled on the same 0.7° grid. Reference datasets are generally highly correlated with the MRM ($R > 0.85$ for all datasets except the coarsest 20CR). This feature is related to the dependence of snow water equivalent on topography, i.e., these datasets represent larger SNW values at higher altitudes. Satellite products and the CFSR reanalysis are very close to the ~~MRM also in terms of NSD and RMSE~~each others, with lower variance with respect to the MRM. The MERRA reanalysis is close to the MRM~~but it has larger~~, with comparable standard deviation and ~~a wider distribution of SNW values, compared to the MRM, satellite products and the CFSR reanalysis~~small RMSE. The ERA-Interim/Land and 20CR reanalyses show opposite behaviors in terms of normalized standard deviation, i.e. very high and very low respectively. ERA-

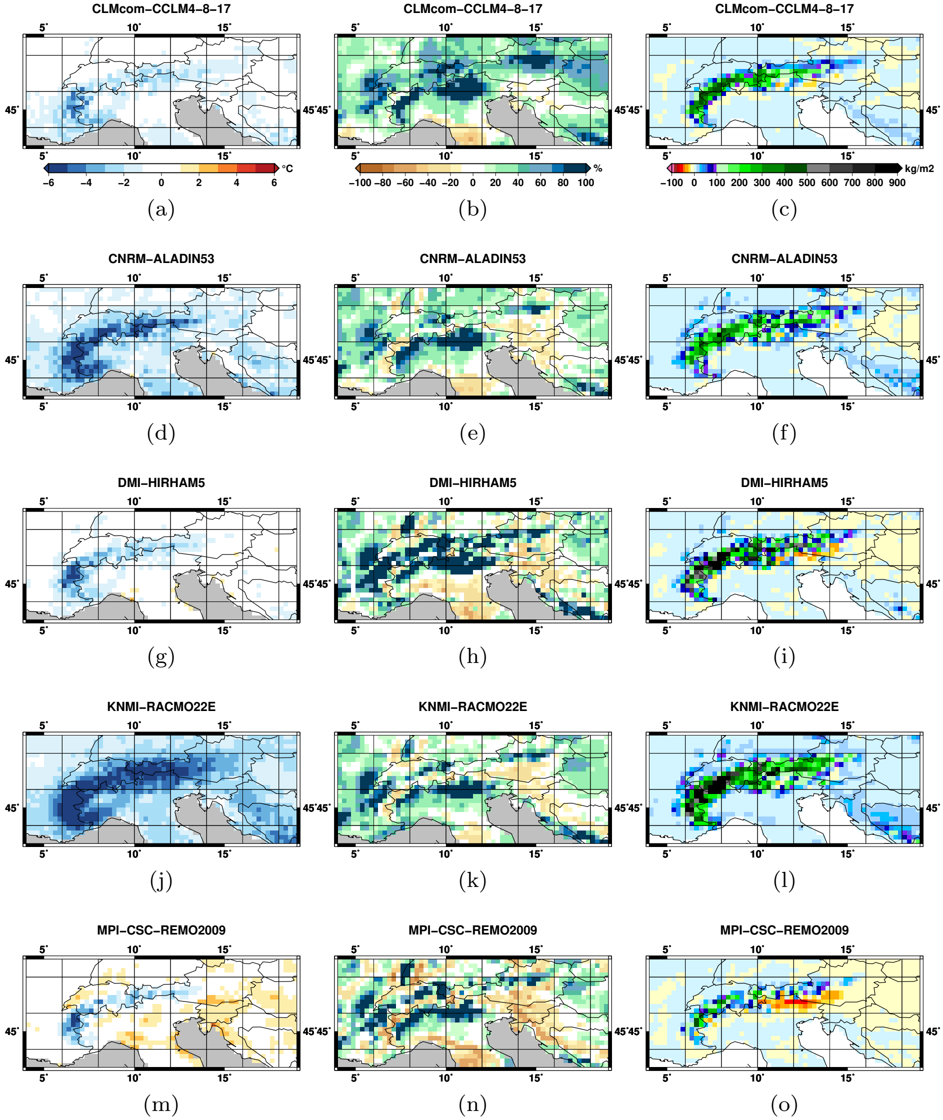
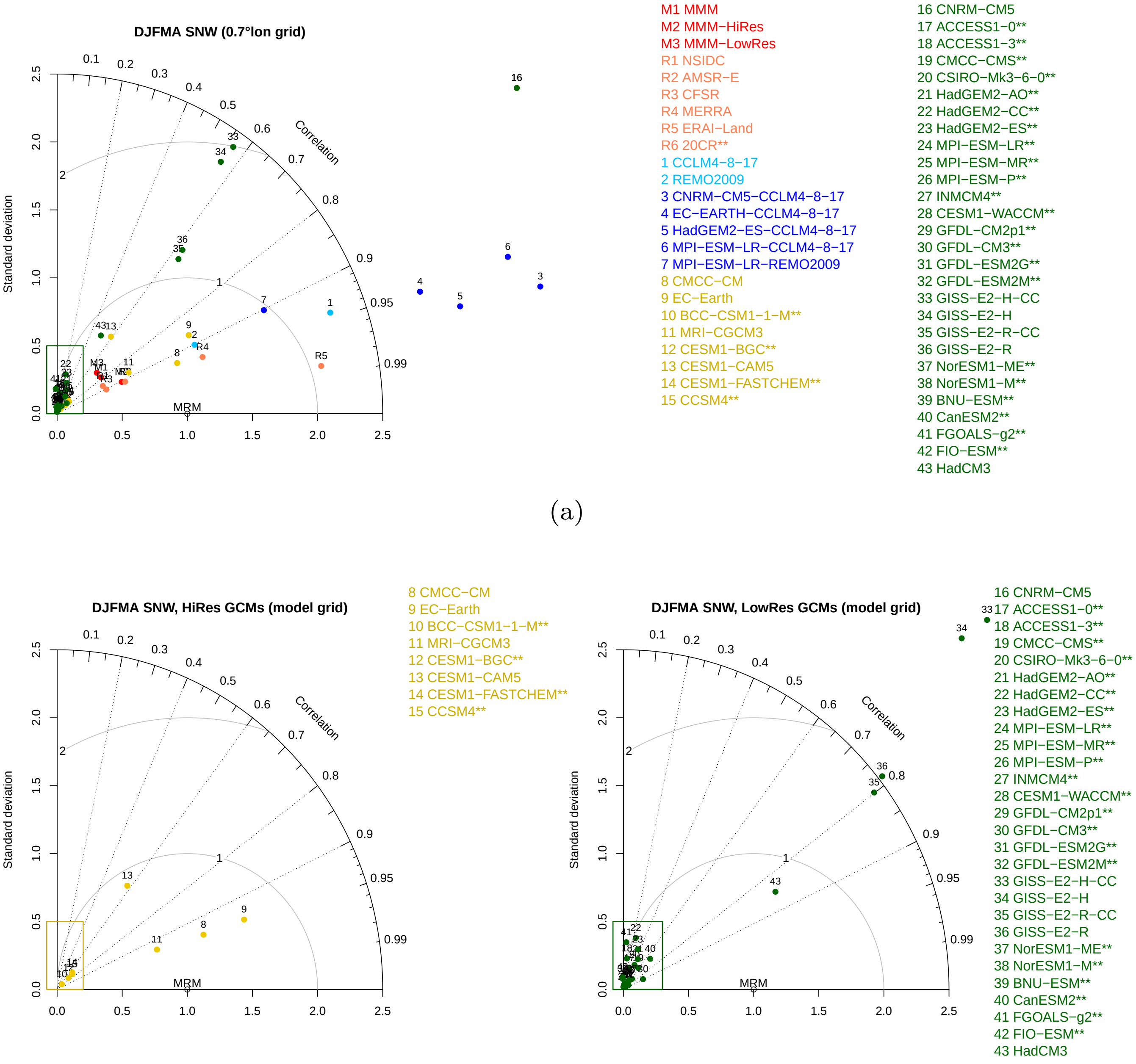


Figure 4. As in ~~Fig.~~ 3 but for ~~global climate models~~the CORDEX ERA-Interim-driven RCM simulations, averaged over the period 1990-2005.



(a)

(b)

Figure 5. Taylor diagrams of the multiannual mean (1980-2005) of the DJFMA average snow water equivalent as described by climate models against the Multi-Reference-Mean (MRM): (a) all datasets are projected onto the same reference grid at 0.7°lon; (b) the climate models are kept at their original resolution and the reference datasets are remapped onto the grid of each model. Points included in the rectangles correspond to models highlighted with ** in the legend.

Interim/Land has a wider ~~distribution~~ statistical dispersion of SNW values and higher SNW peaks, clearly reflected in Fig. 2e, while 20CR has a narrow range of SNW values and a smooth SNW pattern (Fig. 2f).

Of the two RCMs considered, REMO2009 is in better agreement with the MRM in terms of RMSE and NSD. CCLM4-8-17 has large normalized standard deviation, comparable to that found ~~for~~ in ERA-Interim/Land. All GCM-driven simulations show higher variance with respect to the corresponding ERA-Interim driven runs.

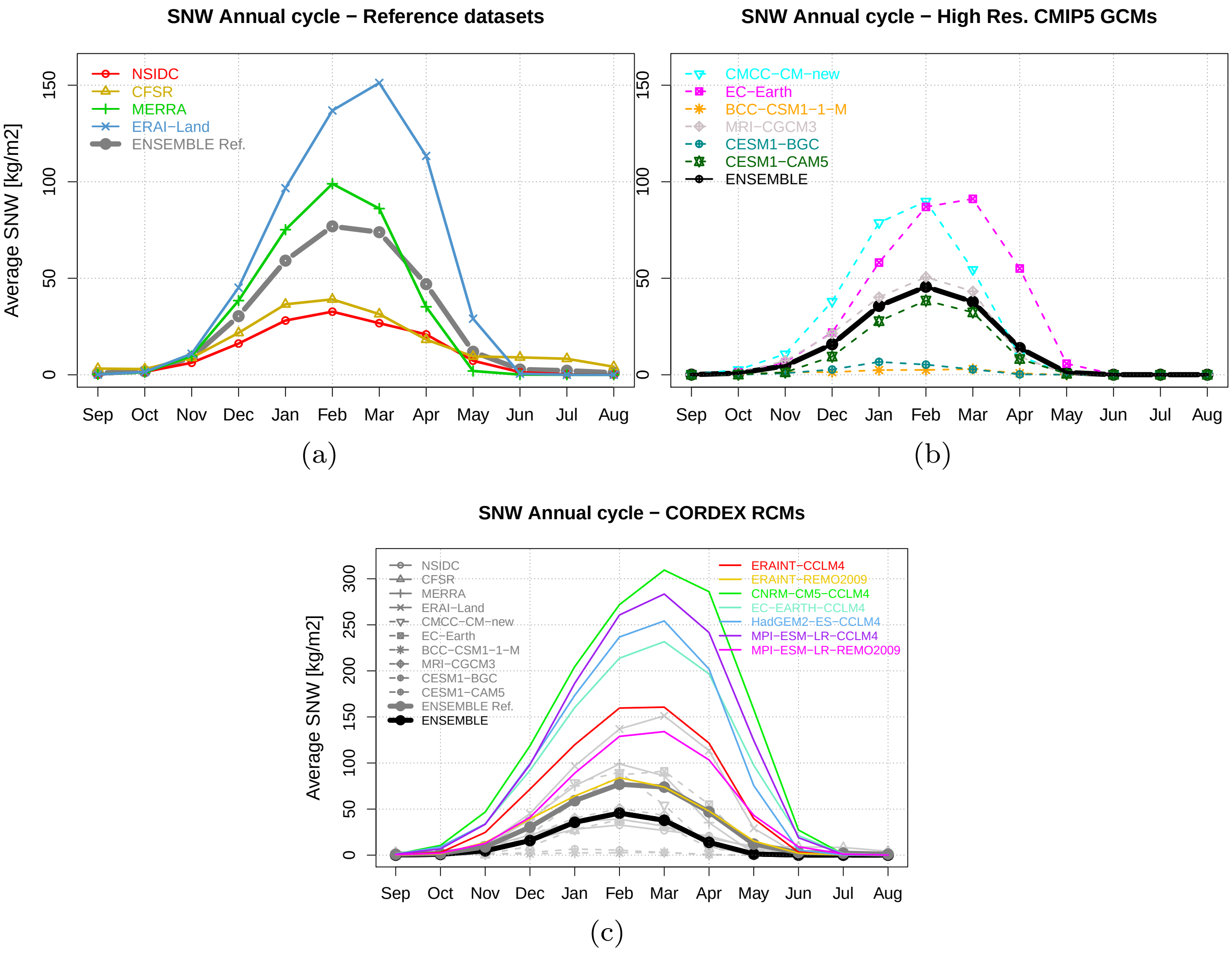
For GCMs, an important feature emerging from this analysis is that, on average, the ensemble mean of the high resolution models performs better in terms of standard deviation, root-mean-square difference and pattern correlation, with respect to the ensemble mean of all CMIP5 GCMs. This result highlights the importance of the horizontal resolution in simulating snowpack spatial patterns (Terzago et al., 2014).

~~To~~ An alternative approach has been devised to provide a fair comparison of the GCMs~~and reduce the impact of the horizontal resolution on their performances, in particular on their spatial variance, each GCM is then~~. Each GCM is compared to the MRM after having ~~remapped each individual~~ conservatively remapped each reference dataset onto the individual ~~GCM~~ GCMs grid, so that the reference is reshaped each time according to the model resolution. This approach allows for a fair ~~comparison also for low resolutionmodels~~evaluation of each GCM on its own grid, regardless of its resolution. For the sake of clarity, we present the results relative to this approach plotting separately the models with resolution *equal or finer* and *coarser* than 1.25° (Fig. 3b). The clustering based on spatial resolution reveals that coarse resolution models generally have very high or very low standard deviation (please note that the CNRM-CM5 model lays outside the range of the plot). In such circumstances the ensemble mean of the models is the result of compensating extreme behaviors, and it should be considered with caution. On the contrary, individual high resolution GCMs are generally closer to the MRM and do not ~~present extreme features: they constitute~~ exhibit extreme features, constituting a more homogeneous ensemble~~that we consider for the subsequent analyses discussed below.~~.

Figure 5 provides information on the similarity of SNW climatologies and, indirectly, qualitative information on the degree of interdependency of the models belonging to the same "family". For example, among the previously mentioned 4 CESM-family models, namely CESM1-CAM5, CESM1-BGC, CESM1-FASTCHEM and CCSM4, three models show a high degree of similarity (Figure 5b). In the calculation of the MMM-HiRes, then, in order to limit the bias related to the interdependency of the models, out of these three similar models we retained only one, CESM1-BGC. In the following, with *high resolution GCMs* we intend 6 models: CMCC-CM, EC-Earth, MRI-CGCM3, BCC-CSM1-1-M, CESM1-BGC, CESM1-CAM5. These models are further analyzed in the following sections. The interdependency of lower resolution GCMs is not clearly detectable from the Taylor diagram and it is not investigated further as these models are not the main focus of the paper, owing to their overall poor performances in the representation of SNW.

4.3 Annual cycle of snow water equivalent

We show in Fig. 6a-b the annual cycle of snow water equivalent as represented by the reference datasets and by the ~~HiRes~~ high resolution GCMs. The monthly SNW at elevation higher than 1000 m a.s.l is spatially averaged over the Greater Alpine Region and temporally averaged over the common period 1980-2005 ~~.~~(see Section 3 for details.).



(a)

(b)

(c)

Figure 6. (a) Annual cycle of snow water equivalent in the reference datasets and (b) in CMIP5 high-resolution GCMs (spatial averages over areas above 1000 m a.s.l., temporal averages over the baseline period 1980-2005). (c) Annual cycle in ERA-Interim-driven and GCM-driven regional climate model simulations, calculated over the period 1990-2005, in comparison to reference datasets and GCM simulations.

The annual cycle in the reference datasets displays a unimodal distribution, with the maximum occurring in different months from Jaunary to March for different datasets. The spread in the reference datasets is quite large, ranging from ~~13~~ about 40 kg/m^2 SNW peak in January in the ~~20CR reanalysis~~ NSIDC satellite product to 150 kg/m^2 SNW peak in March in ERA-Interim/Land. These two products have the most extreme behavior. NSIDC and CFSR show a very similar annual cycle (and comparable spatial patterns), while MERRA presents intermediate values between these two and ERA-Interim/Land. The MRM peaks in February, at about ~~60~~ 75 kg/m^2. The spread among the ~~HiRes GCMs , altough~~ high resolution GCMs is also rather large, ~~is anyway lower than that found~~ as for the reference datasets. Snow water equivalent ~~peaks~~ maximum values range from 3 kg/m^2 according to BCC-CSM1-1-M to about 90 kg/m^2 according to EC-Earth. CESM1-BGC and BCC-CSM-1-1-M show very shallow SNW (few kg/m^2) throughout the year and a much shorter snow season, owing to a large positive bias in air temperature (Fig. 4g,m). CMCC-CM and EC-Earth ~~present~~ display above-average values, with EC-Earth reproducing a snow cycle similar to ERA-Interim/Land ~~, likely because~~ but with lower amplitude. The similarity between EC-Earth and ERA-Interm/Land is likely related to the fact that they use the same land surface model, HTESSEL (Hazeleger et al., 2012). As in the case of the MRM, also the MMM-HiRes peaks in February ~~, with comparable but slightly~~ but with lower SNW values of approximately 50 kg/m^2. With respect to the reference ensemble mean, the GCM ensemble mean tends to underestimate SNW throughout the snow season.

An important outcome of this analysis is that the reference datasets exhibit a large spread in the Alps~~, even larger than that in the high resolution GCMs~~. As a consequence, any ~~assessments~~ assessment based on the use of individual datasets within this ensemble and within this region should be taken with extreme caution.

Figure 6c shows a synthetic view of the SNW annual cycle as in the RCMs simulations compared to the reference datasets and to GCMs. ERA-Interim driven simulations provide similar results ~~as~~ to the reference datasets. In particular the ERA-Interim-REMO2009 annual cycle is close to the ensemble mean of the reference datasets and the ERA-Interim-CCLM4-8-17 annual cycle is close to that provided by ERA-Interim/Land. Relatively larger snow water equivalent values by the CCLM4-8-17 model can be ~~explained since this model was found to have a small cold temperature bias and a wet precipitation bias in the DJFMA the snow season~~ related to wetter conditions (Fig. ~~3~~4a,b) ~~. The combination of colder and wetter conditions may have~~ which probably resulted in larger snow accumulation~~and duration~~.

GCM-driven simulations ~~tend to~~ overestimate the SNW annual cycle in comparison to ~~the~~ their ERA-Interim-driven counterparts. REMO2009, when driven by MPI-ESM-LR GCM, provides SNW values close to the maximum values found in reference datasets, and CCLM-4-8-17, irrespective of the ~~driver~~ driving GCM, shows notably thicker snow pack than any reference datasets and/or GCM. The snow peak is about ~~5~~ 3 times higher than the reference ensemble mean, up to almost twice the ERAInterim-driven value, and it is shifted later in the snow season. Such ~~a result~~ outcome reflects the biases inherent in the driving GCMs, that result in large errors in SNW estimates.

An important hint of this analysis is that despite the large differences in horizontal resolutions, ~~GCMs and~~ the reference datasets, selected high resolution GCMs, and the ERA-Interim-driven ~~RCM~~ RCMs provide comparable results in terms of SNW when the quantities are spatially averaged over the Alpine domain. Unfortunately ~~,~~ the uncertainty on the SNW anual cycle as represented by these datasets is large, and conclusive statements on the accuracy of these SNW estimates ~~, from both RCMs and GCMs,~~ require a reliable ground truth to validate the model results.

4.4 Future changes in the annual cycle of SNW

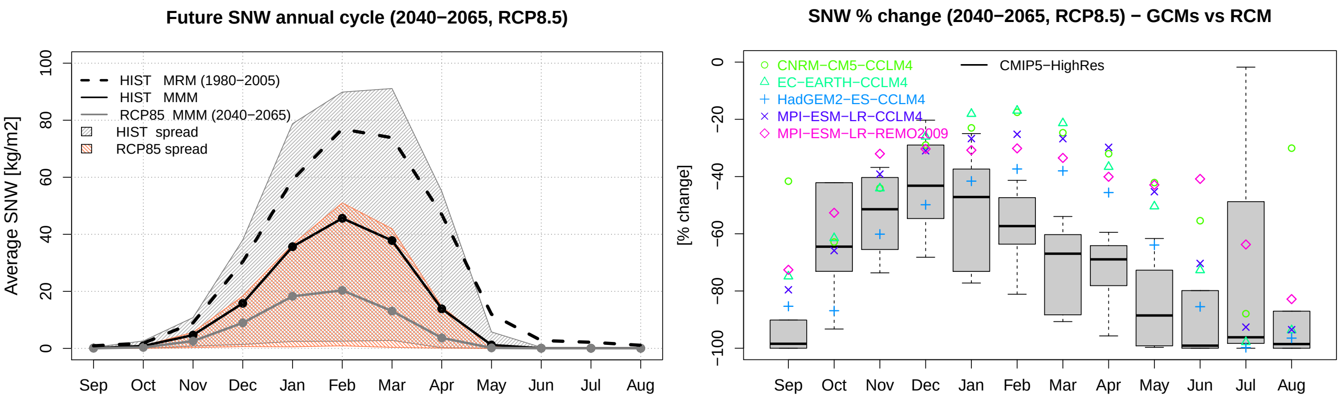
Figure 7a shows the projected annual cycle of snow water equivalent by mid 21^{st} century (2040-2065) in the RCP8.5 scenario compared to the historical annual cycle (1980-2005), according to the high resolution CMIP5 models. Both the ensemble mean and the spread of GCMs are shown. The SNW peak is expected to reduce by more than 50% in the future, with respect to the historical multi-model mean~~, reaching values of about 20 kg/m^2~~. The uncertainty on the amplitude of the snow peak is however very large and the value depends upon the selected GCM. The spread in the percent changes of SNW according to the various models (Fig. 7b) reveals the degree of inter-model consistency. The largest uncertainty is found in summer months, i.e. when snow cover persists only at high altitudes and it can be very shallow. EC-Earth shows a smaller reduction while all the other models predict almost complete snow loss, on average, over the Alpine region (not shown). The lowest reduction is found in December, when the projected decrease ranges between -20% and -70% depending on the model.

For comparison we also analyze the projected changes in SNW annual cycle according to the REMO2009 model and to the CCLM4-8-17 model driven by different GCMs (Fig. 7b). Interestingly, the percent SNW reduction according to RCMs, although still remarkable, is lower compared to CMIP5 GCMs~~.~~, especially in the spring season. From February to April the percent SNW change reported by RCMs lies outside the range of variability of CMIP5 models. The robustness of this result should be verified by considering a larger RCM ensemble, as soon as additional RCM simulations will become available. Fig. 7b ~~shows also~~ also shows the influence of the driving GCM on SNW changes. The spread among the different RCM simulations allows to evaluate the impact of the uncertainty due to the drivers of the snow changes, and its amplitude stresses the importance of performing ensemble analyses.

5 Discussion

We tested the agreement and the uncertainties of the main snow water equivalent datasets – including remote sensing products, reanalyses, global and regional climate models – in reproducing the spatial pattern and the annual cycle of snow over the Greater Alpine Region. The spatial and temporal distribution of SNW is the result of the complex interactions of temperature, precipitation, solar radiation, wind and local geographical features. In mountain areas, in particular, meteo-climatic variables are characterized by high spatial variability depending, among other factors, on elevation, slope, aspect, and exposure to winds. The grid resolution of the remote sensing, reanalysis and climate model products is clearly insufficient to properly represent the spatial variability of snow water equivalent at small scales and at specific locations. For this reason, this study is aimed at analyzing this ensemble of largely used datasets for regional assessment, and quantifying their consistency and degree of agreement in reproducing the average snow conditions at their own resolution.

The reference datasets provide very different pictures of the multiannual mean DJFMA snow water equivalent in the Greater Alpine Region. The satellite-derived datasets and CFSR compare better with each other than with the other products. The



(a) (b)

Figure 7. (a) Annual cycle of snow water equivalent expected by mid 21st century in the RCP8.5 scenario compared to the baseline 1980-2005, as provided by the ~~Hi-Res~~ high resolution CMIP5 models. (b) Percent change in snow water equivalent (2040-2065 average with respect to the baseline 1980-2005) as in the ~~Hi-Res~~ high resolution CMIP5 GCMs (boxplot) and RCM simulations.

two satellite products are based on similar algorithms but rely on different radiometer observations, and AMSR-E doubles the spatial resolution of SMMR and SSM/I. NSIDC and CFSR are likely more similar to each other because CFSR integrates snow analyses based on the same SSM/I observations used by snow algorithm employed in NSIDC (Meng et al., 2012). It is worth stressing that CFSR is, among the reanalyses considered in this study, the only ~~coupled~~ based on atmospheric-ocean-sea ice ~~reanalysis~~coupling; it has the highest horizontal resolution and, as ERA-Interim/Land, it is driven by observed rather than forecast precipitation fields. Interestingly, the analysis system used in CFSR for the atmosphere is similar to the one used in MERRA and despite they use almost the same input data (Saha et al., 2010) they have rather dissimilar snow water equivalent climatologies. MERRA shows a snow distribution comparable to ERA-Interim~~instead~~, likely because they assimilate observations from the same sources and they are run at similar horizontal resolutions. MERRA compares better to the MRM in terms of normalized standard deviation and RMSE, while ERA-Interim ~~presents~~ displays higher snow values in agreement with the results obtained at Northern Hemispheric scale (Mudryk et al., 2015) and over the HKK region (Terzago et al., 2014). The ERA-Interim/Land and 20CR reanalyses show opposite behavior, i.e. very high and very low spatial variability, respectively. In particular the 20CR snow water equivalent fields are extremely smooth with respect to all other datasets. This behavior has been related to a strong warm bias in air temperature in correspondence ~~to~~ of the alpine ridge.

The documented wide range of uncertainty has to be taken into account when using these snow datasets. Some discrepancies can be explained by possible biases in the ~~the~~ drivers of snow processes, the use of different land surface models, different snow schemes and different data assimilation methods, as discussed above. Additional weak points of these products are (i) their low spatial resolution with respect to ~~that which~~ what would be required to represent snowpack processes in mountain environments and (ii) the limited or null constraint by surface snow depth or snow water equivalent observations at high elevations (i.e., no snow assimilation). At global scale, the spread over mountain regions has been estimated to be several times larger than over non-mountainous midlatitude regions (Mudryk et al., 2015). Reducing this gap through improvements in the horizontal resolution and enhanced assimilation of surface data will open new perspectives for a more reliable representation of snow resources in mountain regions at regional to global scale. Efforts have already been spent to provide ~~a reliable atmospheric forcing~~reliable atmospheric fields to land-surface and snow schemes, for example improving precipitation in CFSR and ERA-Interim/Land. Further inclusion of a better resolved topography allows for a more realistic representation of snow processes and could mitigate the issue of upscaling surface measurements at the model grid in the assimilation process.

GCMs have evident limitations in representing the distribution of altitudes in the Greater Alpine region, with the most resolved models underestimating the 95th percentile of the distribution by 500-800 meters. GCMs do not take into proper account ~~the elevations above 1500~~ elevations above 1500-2000 m a.s.l. which are simply non-represented in most models ~~.~~ (see also Fig. S04 for futher details on the elevation ranges represented in each dataset). On the other hand, the analysis of the CMIP5 GCMs reveals that models with spatial resolution finer or equal to 1.25° are in better agreement with the ensemble mean of the reference datasets than the whole GCM ensemble. Compared to low resolution models, the high resolution models form a more homogeneous cluster with no extreme behavior and a higher score (lower RMSE and relative standard deviation closer to one). Provided that high resolution GCMs have different characteristics and different land surface model components (Table 1), their better performance is likely due to the (relatively) finer spatial resolution. This analysis clearly indicates the added value of snow simulations at higher horizontal resolution, even for the typical resolutions of GCMs.

The ~~influence of the single model bias with respect to the reference has been minimized by analyzing the future change in snow water equivalent with respect to the historical mean, i.e. by considering anomalies. GCM projections agree in showing a strong reduction of snow resources by mid-21st century in the RCP 8.5 scenario, on average about 50% in winter and 80% in spring. The uncertainties on the amplitude of the snow water equivalent change are large, but the signal is coherent across all models.~~

~~The~~ EURO-CORDEX regional downscaling experiment further elucidates how the horizontal resolution can affect the representation of the snow processes in mountain areas. The results from the currently available simulations at 0.11° resolution (5 ERA-Interim-driven models) show ~~locally~~ a much thicker average snowpack over the alpine ridge and shallower snowpack at low elevations with respect to the reference dataset. This behavior, related to the RCM finer resolution, is sometimes smothed out when snow water equivalent is spatially averaged over the Alpine domain. At regional scale, the annual cycle represented by ERA-Interim-driven RCMs results comparable to those found in the reference datasets and in GCMs. Important deviations from the reference datasets arise in GCM-driven RCM simulations, owing to the biases inherent in the GCM forcing.

The influence of the single model bias with respect to the reference has been minimized by analyzing the future change in snow water equivalent with respect to the historical mean, i.e. by considering anomalies. GCM projections agree in showing a strong reduction of snow resources by mid-21st century in the RCP 8.5 scenario, especially in the spring season. The uncertainties on the amplitude of the snow water equivalent change are large, but the signal is coherent across all models.

RCMs future projections show weaker snow reductions with respect to the ~~coarse~~ coarser scale high resolution GCMs, especially in the spring season, when future snow projections appear particularly uncertain. While few ~~RCM~~ regional models can have limited representativess of the whole EURO-CORDEX ensemble and a larger set of simulations has to be considered as soon as they become available, this analysis highlights the large discrepancy among the considered datasets over the historical period and calls for a reference observation-based product that could reliably represent the ground truth.

6 Conclusions

This study shows that the spatial and temporal distribution of snow water equivalent in the Greater Alpine Region (one of the most measured mountain regions in the world) is quite uncertain. The ~~major~~ main available gridded snow water equivalent datasets are derived from remote sensing observations and reanalyses but they have never been properly ~~evaluated~~ validated in mountain regions owing to the limited availability of in situ snow observations. In this work, we compared such datasets to highlight the degree of agreement in the mean climatologies, to quantify their spread and assess the uncertainties associated ~~to~~ with snow estimates. These datasets provide very different pictures of the snow spatial distribution and seasonal cycle. Of course, mountain regions are non-optimal conditions to test these coarse-grid datasets, as surface heterogeneity at ~~the~~ sub-grid scale is difficult to represent for both remote sensing and reanalysis data. This argument enforces the evidence that we currently lack proper information on snowpack distribution at mountain range scale. Knowledge of the long-term variability of the snowpack at *high spatial resolution* and at *mountain range* scale is limited but ~~dramatically~~ necessary for climate studies, for calibrating/validating models, for data assimilation in the ~~reanalyses~~ reanalysis products and for assessing seasonal water resources. In our opinion, improving the open availability and the exchange of in-situ snow observations and developing gridded snow datasets representative of the ground truth in mountain regions is a highest priority for advancing cryospheric/hydrologic research in mountain environments.

A second action for improving snow estimates in mountain areas in both reanalyses and climate models is to pursue high resolution simulations, to allow for a better representation of the main drivers of the snow processes, i.e. temperature and precipitation patterns and their dependence on elevation. An increased horizontal resolution, and thus a more accurate representation of topography, allows for a better description of the spatial distribution and phase of precipitation and of altitudinal temperature gradients. New insights on this topic are expected by the High ~~Resolution~~ RESolution Model Intercomparison Project (Haarsma et al., 2016), the CMIP6-endorsed coordinated experiment that will provide an ensemble of GCM runs at spatial resolutions significantly ~~higher~~ finer than the current generation CMIP5 models.

A further goal is the refinement of the representation of snowpack processes, that at the moment are drastically simplified, in global climate and earth system models (ESMs). This issue is being addressed by the ESM-SnowMIP initiative (van den Hurk et al., 2016, see also http://www.climate-cryosphere.org/activities/targeted/esm-snowmip) through coordinated experiments to evaluate snow modules of large-scale climate models and quantify the required complexity to be represented in ESMs.

The present study contributes to these main challenges by providing a picture of the main available snow products and measuring the related uncertainties in the Alpine environment. The relative assessment of the capability of satellite-based products, reanalyses, RCMs and GCMs in reproducing snowpack features provides important information to both model developers and to the community of users, allowing to identify criticalities in the model components and to be aware of the strengths and limits of the available products.

7 ~~ERA-Interim/Land precipitation and SNW compared to ERA-Interim~~

~~ERA-Interim/Land and ERA-Interim snow water equivalent climatologies are derived using the ECMWF land surface model HTESSEL, being ERA-Interim/Land the result of *offline* simulation driven by meteorological forcing from the ERA-Interim atmospheric reanalysis and precipitation adjustments based on GPCP v2.1.~~

~~Percent differences of DJFMA precipitation forcing in ERA-Interim/Land (Fig. 2n) with respect to ERA-Interim in the Alpine region are reported in Fig. **??**a. ERA-Interim/Land presents a larger precipitation amount over the Alpine range, partially compensating the original ERA-Interim dry bias. The additional precipitation input is reflected in a thicker snowpack, locally exceeding ERA-Interim values by more than 100 kg/m^2.~~

~~a) Percent difference in the multiannual mean (1980-2005) of the DJFMA accumulated precipitation in ERA-Interim/Land with respect to ERA-Interim; (b) Bias of ERA-Interim/Land DJFMA average snow water equivalent climatolology (1980-2005) with respect to ERA-Interim.~~

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreements No. 641816 (CRESCENDO) and No. 641762 (ECOPOTENTIAL). This work was also supported by the Italian project of Interest NextData of the Italian Ministry for Education, University and Research.

We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling and Working Group on Regional Climate, which are responsible for CMIP5 and CORDEX, and we thank the climate modeling groups (listed in Tables 1 and 2) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

References

Armstrong, R., Brodzik, M., Knowles, K., and Savoie, M.: Global Monthly EASE-Grid Snow Water Equivalent Climatology [1979-2005], 2005.

Arora, V., Scinocca, J., Boer, G., Christian, J., Denman, K., Flato, G., Kharin, V., Lee, W., and Merryfield, W.: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, Geophys. Res. Lett., 38, doi:10.1029/2010gl046270, 2011.

Auer, I., Böhm, R., Jurkovic, A., Lipa, W., Orlik, A., Potzmann, R., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., et al.: HISTALP–historical instrumental climatological surface time series of the Greater Alpine Region, International Journal of Climatology, 27, 17–46, 2007.

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., and Betts, A. K.: A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System, J. Hydrometeor., 10, 623–643, doi:10.1175/2008jhm1068.1, 2009.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., et al.: ERA-Interim/Land: a global land water resources dataset., Hydrol. Earth Syst. Sci. Discuss., 10, 14 705–14 745, doi:10.5194/hessd-10-14705-2013, 2013.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., et al.: ERA-Interim/Land: a global land surface reanalysis data set, Hydrology and Earth System Sciences, 19, 389–407, 2015.

Barnett, T. P., Adam, J. C., and Lettenmaier, D. P.: Potential impacts of a warming climate on water availability in snow-dominated regions, Nature, 438, 303–309, 2005.

Beniston, M.: Climatic change in mountain regions: a review of possible impacts, in: Climate variability and change in high elevation regions: Past, present & future, pp. 5–31, Springer, 2003.

Beniston, M. and Stoffel, M.: Assessing the impacts of climatic change on mountain water resources, Science of the Total Environment, 493, 1129–1137, 2014.

Beniston, M., Uhlmann, B., Goyette, S., and Lopez-Moreno, J. I.: Will snow-abundant winters still exist in the Swiss Alps in an enhanced greenhouse climate?, International Journal of Climatology, 31, 1257–1263, 2011.

Bentsen, M., Bethke, I., Debernard, J., Iversen, T., Kirkevåg, A., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I., Hoose, C., et al.: The Norwegian Earth System Model, NorESM1-M–Part 1: Description and basic evaluation of the physical climate, Geosci. Model Dev., 6, 687–720, doi:10.5194/gmd-6-687-2013, 2013.

Bi, D., Dix, M., Marsl, S., O'Farrell, S., Rashid, H., Uotila, P., Hirst, A., Kowalczyk, E., Golebiewski, M., Sullivan, A., et al.: The ACCESS coupled model: Description, control climate and evaluation, Aust. Met. Oceanog. J., 63, 9–32, 2013.

Bøssing Christensen, O., Drews, M., Hesselbjerg Christensen, J., Dethloff, K., Ketelsen, K., Hebestadt, I., and Rinke, A.: The HIRHAM Regional Climate Model. Version 5 (beta), Tech. rep., Danish Climate Centre, Danish Meteorological Institute, 2007.

Brown, R., Derksen, C., and Wang, L.: A multi-data set analysis of variability and change in Arctic spring snow cover extent, 1967–2008, Journal of Geophysical Research: Atmospheres, 115, 2010.

Brun, E., Vionnet, V., Boone, A., Decharme, B., Peings, Y., Valette, R., Karbou, F., and Morin, S.: Simulation of northern eurasian local snow depth, mass, and density using a detailed snowpack model and meteorological reanalyses, Journal of Hydrometeorology, 14, 203–219, 2013.

CDO: CDO 2015: Climate Data Operators, Availableat:http://www.mpimet.mpg.de/cdo, 2015.

Chimani, B., Böhm, R., Matulla, C., and Ganekind, M.: Development of a longterm dataset of solid/liquid precipitation, Advances in Science and Research, 6, 39–43, doi:10.5194/asr-6-39-2011, http://www.adv-sci-res.net/6/39/2011/, 2011.

Clifford, D.: Global estimates of snow water equivalent from passive microwave instruments: history, challenges and future developments, International Journal of Remote Sensing, 31, 3707–3726, 2010.

Collier, M., Jeffrey, S., Rotstayn, L., Wong, K., Dravitzki, S., Moseneder, C., Hamalainen, C., Syktus, J., Suppiah, R., Antony, J., et al.: The CSIRO-Mk3. 6.0 Atmosphere-Ocean GCM: participation in CMIP5 and data publication, in: International Congress on Modelling and Simulation-MODSIM 2011, 2011.

Collins, M., Tett, S., and Cooper, C.: The internal climate variability of HadCM3, a version of the Hadley Centre coupled model without flux adjustments, Climate Dynamics, 17, 61–81, 2001.

Collins, W., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C., Joshi, M., Liddicoat, S., et al.: Development and evaluation of an Earth-system model–HadGEM2, Geosci. Model Dev., 4, 1051–1075, doi:10.5194/gmd-4-1051-2011, 2011.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, O., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The Twentieth Century Reanalysis Project, Q. J. Roy. Meteor. Soc., 137, 1–28, doi:10.1002/qj.776, 2011.

Davini, P., Hardenberg, J. v., Corti, S., Christensen, H. M., Juricke, S., Subramanian, A., Watson, P. A., Weisheimer, A., and Palmer, T. N.: Climate SPHINX: evaluating the impact of resolution and stochastic physics parameterisations in the EC-Earth global climate model, Geoscientific Model Development, 10, 1383–1402, 2017.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, doi:10.1002/qj.828, 2011.

Delworth, T. L., Broccoli, A. J., Rosati, A., Stouffer, R. J., Balaji, V., Beesley, J. A., Cooke, W. F., Dixon, K. W., Dunne, J., Dunne, K., et al.: GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics, Journal of Climate, 19, 643–674, 2006.

Diffenbaugh, N. S., Scherer, M., and Ashfaq, M.: Response of snow-dependent hydrologic extremes to continued global warming, Nature climate change, 3, 379–384, 2013.

Donner, L. J., Wyman, B. L., Hemler, R. S., Horowitz, L. W., Ming, Y., Zhao, M., Golaz, J.-C., Ginoux, P., Lin, S.-J., Schwarzkopf, M. D., et al.: The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3, Journal of Climate, 24, 3484–3519, 2011.

Dunne, J. P., John, J. G., Adcroft, A. J., Griffies, S. M., Hallberg, R. W., Shevliakova, E., Stouffer, R. J., Cooke, W., Dunne, K. A., Harrison, M. J., et al.: GFDL's ESM2 global coupled climate-carbon Earth System Models. Part I: Physical formulation and baseline simulation characteristics, Journal of Climate, 25, 6646–6665, 2012.

Dutra, E., Balsamo, G., Viterbo, P., Miranda, P. M., Beljaars, A., Schär, C., and Elder, K.: An improved snow scheme for the ECMWF land surface model: description and offline validation, J. Hydrometeor., 11, 899–916, doi:10.1175/2010jhm1249.1, 2010.

Farda, A., Déué, M., Somot, S., Horányi, A., Spiridonov, V., and Tóth, H.: Model ALADIN as regional climate model for Central and Eastern Europe, Studia Geophysica et geodaetica, 54, 313–332, 2010.

Gan, T. Y., Barry, R. G., and Gobena, A. K.: Changes in Snowpacks of Canadian Prairies for 1979–2004 Detected from Snow Water Equivalent Data of SMMR and SSM/I Passive Microwave and Related Climatic Factors, Remote Sensing of the Terrestrial Water Cycle, 206, 227, 2014.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., et al.: The Community Climate System Model Version 4., J. Climate, 24, doi:10.1175/2011jcli4083.1, 2011.

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W., Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K. D., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K.-H., Claussen, M., Marotzke, J., and Stevens, B.: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5, J. Adv. Model. Earth Syst., 5, 572–597, doi:10.1002/jame.20038, 2013.

Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F., and Wood, R. A.: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments, Climate dynamics, 16, 147–168, 2000.

Haarsma, R. J., Roberts, M., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L. R., Lu, J., Luo, J.-J., Mao, J., Mizielinski, M. S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., and von Storch, J.-S.: High Resolution Model Intercomparison Project (HighResMIP), Geoscientific Model Development Discussions, 2016, 1–35, doi:10.5194/gmd-2016-66, http://www.geosci-model-dev-discuss.net/gmd-2016-66/, 2016.

Hagemann, S.: An improved land surface parameter dataset for global and regional climate models, 2002.

Hall, D. K. and Riggs, G. A.: Accuracy assessment of the MODIS snow products, Hydrological Processes, 21, 1534–1547, 2007.

Hancock, S., Baxter, R., Evans, J., and Huntley, B.: Evaluating global snow water equivalent products for testing land surface models, Remote Sensing of Environment, 128, 107–117, 2013.

Hantel, M., Maurer, C., and Mayer, D.: The snowline climate of the Alps 1961–2010, Theoretical and Applied Climatology, 110, 517–537, 2012.

Hastings, D. and Dunbar, P.: Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, Documentation, Volume 1.0. Key to Geophysical Records Documentation (KGRD) 34., 1999.

Haylock, M., Hofstra, N., Klein Tank, A., Klok, E., Jones, P., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, Journal of Geophysical Research: Atmospheres, 113, 2008.

Hazeleger, W., Wang, X., Severijns, C., Ştefănescu, S., Bintanja, R., Sterl, A., Wyser, K., Semmler, T., Yang, S., and Van den Hurk, B.: EC-Earth V2. 2: description and validation of a new seamless earth system prediction model, Climate Dyn., 39, 2611–2629, doi:10.1007/s00382-011-1228-5, 2012.

Hurrell, J. W., Holland, M., Gent, P., Ghan, S., Kay, J. E., Kushner, P., Lamarque, J.-F., Large, W., Lawrence, D., Lindsay, K., et al.: The Community Earth System Model: A Framework for Collaborative Research, Bulletin of the American Meteorological Society, 94, doi:10.1175/BAMS-D-12-00121.1, 2013.

Imperio, S., Bionda, R., Viterbi, R., and Provenzale, A.: Climate change and human disturbance can lead to local extinction of Alpine rock ptarmigan: New insight from the Western Italian Alps, PloS one, 8, e81 598, 2013.

IPCC: Climate Change 2013: The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2013.

Jacob, D. and Podzun, R.: Sensitivity studies with the regional climate model REMO, Meteorology and Atmospheric Physics, 63, 119–129, 1997.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., et al.: EURO-CORDEX: new high-resolution climate change projections for European impact research, Regional Environmental Change, 14, 563–578, 2014.

Johns, T., Gregory, J., Ingram, W., Johnson, C., Jones, A., Lowe, J., Mitchell, J., Roberts, D., Sexton, D., Stevenson, D., Tett, S., and Woodage, M.: Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emissions scenarios, Climate Dynamics, 20, 583–612, doi:10.1007/s00382-002-0296-y, 2003.

Jones, P. W.: First-and second-order conservative remapping schemes for grids in spherical coordinates, Monthly Weather Review, 127, 2204–2210, 1999.

Kopp, T. and Kiess, R.: The air force global weather central snow analysis model, in: CONFERENCE ON WEATHER ANALYSIS AND FORECASTING, vol. 15, pp. 220–222, American Meteorological Society, 1996.

Körner, C.: Alpine plant life: functional plant ecology of high mountain ecosystems, Springer Science & Business Media, 2003.

Koster, R. D., Suarez, M. J., Ducharne, A., Stieglitz, M., and Kumar, P.: A Catchment-Based Approach to Modeling Land Surface Processes in a GCM. Part 1; Model Structure, Journal Geophysical Research, 105, 24 809–24 822, 2000.

Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E., et al.: Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble, Geoscientific Model Development, 7, 1297–1333, 2014.

Li, L., Lin, P., Yu, Y., Wang, B., Zhou, T., Liu, L., Liu, J., Bao, Q., Xu, S., Huang, W., Xia, K., Pu, Y., Dong, L., Shen, S., Liu, Y., Hu, N., Liu, M., Sun, W., Shi, X., Zheng, W., Wu, B., Song, M., Liu, H., Zhang, X., Wu, G., Xue, W., Huang, X., Yang, G., Song, Z., and Qiao, F.: The flexible global ocean-atmosphere-land system model, Grid-point Version 2: FGOALS-g2, Advances in Atmospheric Sciences, 30, 543–560, doi:10.1007/s00376-012-2140-6, 2013.

Meng, J., Yang, R., Wei, H., Ek, M., Gayno, G., Xie, P., and Mitchell, K.: The land surface analysis in the NCEP Climate Forecast System Reanalysis, Journal of Hydrometeorology, 13, 1621–1630, 2012.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. A., Mitchell, J. F. B., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P., and Wilbanks, T. J.: The next generation of scenarios for climate change research and assessment, Nature, 463, 747–756, doi:10.1038/nature08823, 2010.

Mudryk, L., Derksen, C., Kushner, P., and Brown, R.: Characterization of Northern Hemisphere snow water equivalent datasets, 1981–2010, Journal of Climate, 28, 8037–8051, 2015.

Qiao, F., Song, Z., Bao, Y., Song, Y., Shu, Q., Huang, C., and Zhao, W.: Development and evaluation of an Earth System Model with surface gravity waves, J. Geophys. Res. Oceans, 118, doi:10.1002/jgrc.20327, 2013.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., et al.: MERRA: NASA's modern-era retrospective analysis for research and applications, Journal of Climate, 24, 3624–3648, 2011.

Rixen, C., Teich, M., Lardelli, C., Gallati, D., Pohl, M., Pütz, M., and Bebi, P.: Winter tourism and climate change in the Alps: an assessment of resource consumption, snow reliability, and future snowmaking potential, Mountain Research and Development, 31, 229–236, 2011.

Rockel, B., Will, A., and Hense, A.: The regional climate model COSMO-CLM (CCLM), Meteorologische Zeitschrift, 17, 347–348, 2008.

Sabin, T., Krishnan, R., Ghattas, J., Denvil, S., Dufresne, J.-L., Hourdin, F., and Pascal, T.: High resolution simulation of the South Asian monsoon using a variable resolution global climate model, Climate dynamics, 41, 173–194, 2013.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., et al.: The NCEP Climate Forecast System Reanalysis, B. Am. Meteorol. Soc, 91, 1015–1057, doi:10.1175/2010BAMS3001.1, 2010.

Salzmann, N., Huggel, C., Rohrer, M., and Stoffel, M.: Data and knowledge gaps in glacier, snow and related runoff research–A climate change adaptation perspective, Journal of Hydrology, 518, 225–234, 2014.

Scherrer, S., Ceppi, P., Croci-Maspoli, M., and Appenzeller, C.: Snow-albedo feedback and Swiss spring temperature trends, Theoretical and Applied Climatology, 110, 509–516, 2012.

Schmidt, G. A., Ruedy, R., Hansen, J. E., Aleinov, I., Bell, N., Bauer, M., Bauer, S., Cairns, B., Canuto, V., Cheng, Y., et al.: Present-day atmospheric simulations using GISS ModelE: Comparison to in situ, satellite, and reanalysis data., Journal of Climate, 19, 2006.

Scoccimarro, E., Gualdi, S., Bellucci, A., Sanna, A., Giuseppe Fogli, P., Manzini, E., Vichi, M., Oddo, P., and Navarra, A.: Effects of tropical cyclones on ocean heat transport in a high-resolution coupled general circulation model, J. Climate, 24, 4368–4384, doi:10.1175/2011jcli4104.1, 2011.

Serquet, G., Marty, C., Dulex, J.-P., and Rebetez, M.: Seasonal trends and temperature dependence of the snowfall/precipitation-day ratio in Switzerland, Geophysical research letters, 38, 2011.

Steger, C., Kotlarski, S., Jonas, T., and Schär, C.: Alpine snow cover in a changing climate: a regional climate model perspective, Climate dynamics, 41, 735–754, 2013.

Stewart, I. T.: Changes in snowpack and snowmelt runoff for key mountain regions, Hydrological Processes, 23, 78–94, 2009.

Stieglitz, M., Ducharne, A., Koster, R., and Suarez, M.: The impact of detailed snow physics on the simulation of snow cover and subsurface thermodynamics at continental scales, Journal of Hydrometeorology, 2, 228–242, 2001.

Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Kärnä, J.-P., Koskinen, J., and Bojkov, B.: Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements, Remote Sensing of Environment, 115, 3517–3529, 2011.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, Journal of Geophysical Research: Atmospheres, 106, 7183–7192, 2001.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American Meteorological Society, 93, 485, 2012.

Tedesco, M., Kelly, R., Foster, J., and Chang, A.: AMSR-E/Aqua Daily L3 Global Snow Water Equivalent EASE-Grids. Version 2. [2002-2011], 2004.

Terzago, S., von Hardenberg, J., Palazzi, E., and Provenzale, A.: Snowpack changes in the Hindu-Kush Karakoram Himalaya from CMIP5 Global Climate Models, Journal of Hydrometeorology, 15, 2014.

Tong, J. and Velicogna, I.: A comparison of AMSR-E/Aqua snow products with in situ observations and MODIS snow cover products in the Mackenzie River Basin, Canada, Remote Sensing, 2, 2313–2322, 2010.

Uppala, S. M., Kållberg, P., Simmons, A., Andrae, U., Bechtold, V. d., Fiorino, M., Gibson, J., Haseler, J., Hernandez, A., Kelly, G., et al.: The ERA-40 re-analysis, Quarterly Journal of the Royal Meteorological Society, 131, 2961–3012, 2005.

van den Hurk, B., Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., Douville, H., Colin, J., Ducharne, A., Cheruy, F., Viovy, N., Puma, M. J., Wada, Y., Li, W., Jia, B., Alessandri, A., Lawrence, D. M., Weedon, G. P., Ellis, R., Hagemann, S., Mao, J., Flanner, M. G., Zampieri, M., Materia, S., Law, R. M., and Sheffield, J.: LS3MIP (v1.0) contribution to CMIP6: the Land Surface, Snow and Soil moisture Model Intercomparison Project – aims, setup and expected outcome, Geoscientific Model Development, 9, 2809–2832, doi:10.5194/gmd-9-2809-2016, http://www.geosci-model-dev.net/9/2809/2016/, 2016.

Van Meijgaard, E., Van Ulft, L., Lenderink, G., De Roode, S., Wipfler, E. L., Boers, R., and van Timmermans, R.: Refinement and application of a regional atmospheric model for climate scenario calculations of Western Europe, KVR 054/12, KVR, 2012.

Voldoire, A., Sanchez-Gomez, E., Salas y Mélia, D., Decharme, B., Cassou, C., Sénési, S., Valcke, S., Beau, I., Alias, A., Chevallier, M., Déqué, M., Deshayes, J., Douville, H., Fernandez, E., Madec, G., Maisonnave, E., Moine, M.-P., Planton, S., Saint-Martin, D., Szopa, S., Tyteca, S., Alkama, R., Belamari, S., Braun, A., Coquart, L., and Chauvin, F.: The CNRM-CM5.1 global climate model: description and basic evaluation, Climate Dyn., 40, 2091–2121, doi:10.1007/s00382-011-1259-y, 2013.

Volodin, E., Dianskii, N., and Gusev, A.: Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations, Atmos. Oceanic Phys., 46, 414–431, doi:10.1134/s000143381004002x, 2010.

Wu, T., Li, W., Ji, J., Xin, X., Li, L., Wang, Z., Zhang, Y., Li, J., Zhang, F., Wei, M., et al.: Global carbon budgets simulated by the Beijing Climate Center Climate System Model for the last century, J. Geophys. Res. Atmos., 118, 4326–4347, doi:10.1002/jgrd.50320, 2013.

Yukimoto, S., Adachi, Y., and Hosaka, M.: A new global climate model of the Meteorological Research Institute: MRI-CGCM3: model description and basic performance (special issue on recent development on climate models and future climate projections), J. Meteorol. Soc. Jpn, 90, 23–64, doi:10.2151/jmsj.2012-a02, 2012.

Table 1. Snow water equivalent datasets, including remote sensing products, reanalyses and CMIP5 Global Climate Models, used in this study. For each of these we report the land surface model (LSM, when it applies), the spatial/spectral horizontal resolution and the relevant references. CMIP5 models with horizontal resolution equal or finer than 1.25° ~~lon~~ longitude are highlighted in bold.

Model	Institution	LSM	Res. [°lon]/Sp.Res	Reference
Global SWE	National Snow and Ice Data Center	–	25 km	Armstrong et al. (2005)
AMSR-E/Aqua Monthly L3 Global SWE	National Snow and Ice Data Center	–	25 km	Tedesco et al. (2004)
CFSR	US National Centers for Environmental Prediction	Noah	0.3125	Saha et al. (2010)
MERRA	US National Aeronautics and Space Administration	Catchment LSM	0.67	Rienecker et al. (2011)
ERA-Interim/Land	European Centre for Medium-Range Weather Forecasts	HTESSEL	0.7	Balsamo et al. (2013)
20th Century Reanalysis	NOAA Earth System Research Laboratory	Noah	1.875	Compo et al. (2011)
CMCC-CM	Euro-Mediterranean Centre for Climate Change	ECHAM5	**0.75** / T159	Scoccimarro et al. (2011)
EC-Earth	EC-Earth Consortium	HTESSEL	**1.125** / T159	Hazeleger et al. (2012)
BCC-CSM1.1m	Beijing Climate Center, China	BCC_AVIM1.0	**1.125** / T106	Wu et al. (2013)
MRI-CGCM3	Meteorological Research Institute, Japan	HAL	**1.125** / T159	Yukimoto et al. (2012)
CESM1-BGC	National Center for Atmospheric Research	CLM4	**1.25**	Hurrell et al. (2013)
CESM1-CAM5	National Center for Atmospheric Research	CLM4	**1.25**	Hurrell et al. (2013)
CESM1-FASTCHEM	National Center for Atmospheric Research	CLM4	1.25	Hurrell et al. (2013)
CCSM4	National Center for Atmospheric Research	CLM4	1.25	Gent et al. (2011)
CNRM-CM5	Centre National de Recherches Météorologiques	ISBA	1.4 / T127	Voldoire et al. (2013)
ACCESS1-0	CSIRO/BOM, Australia	MOSES2	1.875 / N96	Bi et al. (2013)
ACCESS1-3	CSIRO/BOM, Australia	CABLE1.0	1.875 / N96	Bi et al. (2013)
CMCC-CMS	Euro-Mediterranean Centre for Climate Change	ECHAM5	1.875 / T63	Scoccimarro et al. (2011)
CSIRO-Mk3-6-0	CSIRO, Australia	MOSES II	1.875 / T63	Collier et al. (2011)
HadGEM2-AO	Met Office Hadley Centre	MOSES II	1.875 / N96	Collins et al. (2011)
HadGEM2-CC	Met Office Hadley Centre	MOSES II	1.875 / N96	Collins et al. (2011)
HadGEM2-ES	Met Office Hadley Centre	MOSES II	1.875 / N96	Collins et al. (2011)
MPI-ESM-LR	Max Planck Institute for Meteorology	JSBACH	1.875 / T63	Giorgetta et al. (2013)
MPI-ESM-MR	Max Planck Institute for Meteorology	JSBACH	1.875 / T63	Giorgetta et al. (2013)
MPI-ESM-P	Max Planck Institute for Meteorology	JSBACH	1.875 / T63	Giorgetta et al. (2013)
INM-CM4	Institute for Numerical Mathematics	INM	2.0	Volodin et al. (2010)
CESM1-WACCM	National Center for Atmospheric Research	CAM	2.5	Hurrell et al. (2013)
GFDL-CM3	NOAA Geophysical Fluid Dynamics Laboratory	LM3	2.5	Donner et al. (2011)
GFDL-ESM2G	NOAA Geophysical Fluid Dynamics Laboratory	LM3	2.5	Dunne et al. (2012)
GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory	LM3	2.5	Dunne et al. (2012)
GFDL-CM2p1	NOAA Geophysical Fluid Dynamics Laboratory	LM2	2.5	Delworth et al. (2006)
GISS-E2-H-CC	NASA Goddard Institute for Space Studies	GISS LSM	2.5	Schmidt et al. (2006)
GISS-E2-H	NASA Goddard Institute for Space Studies	GISS LSM	2.5	Schmidt et al. (2006)
GISS-E2-R-CC	NASA Goddard Institute for Space Studies	GISS LSM	2.5	Schmidt et al. (2006)
GISS-E2-R	NASA Goddard Institute for Space Studies	GISS LSM	2.5	Schmidt et al. (2006)
NorESM1-ME	Norwegian Climate Centre	CLM4	2.5	Bentsen et al. (2013)
NorESM1-M	Norwegian Climate Centre	CLM4	2.5	Bentsen et al. (2013)
BNU-ESM	Beijing Normal University, China	BNU-CoLM3	2.8125 / T42	[1]
CanESM2	Canadian Centre for Climate Modelling and Analysis	CLASS	2.8125 / T63	Arora et al. (2011)
FGOALS-g2	LASG/CESS, China	CLM3	2.8125	Li et al. (2013)
FIO-ESM	The First Institute of Oceanography, China	CLM3.5	2.8125 / T42	Qiao et al. (2013)
HadCM3	Met Office Hadley Centre	MOSES I	3.75 / N48	Johns et al. (2003)

Reference: [1]=http://esg.bnu.edu.cn/BNU_ESM_webs/htmls/index.html.

Table 2. EURO-CORDEX Regional Climate Models providing ERA-Interim driven runs for the snow water equivalent variable at 0.11° spatial resolution considered in this study. For each of model we report also the land surface model (LSM), the number of available GCM-driven runs and the ~~relevant references~~reference.

Model	Institution	LSM	Ensemble members	Reference
CCLM4-8-17	CLM Community	Terra-ML	4	Rockel et al. (2008)
ALADIN53	Centre National de Recherches Météorologiques	ISBA	-	Farda et al. (2010)
HIRHAM5	Danish Meteorological Institute	Hagemann (2002)	1	Bøssing Christensen et al. (2007)
RACMO22E	Royal Netherlands Meteorological Institute	HTESSEL	2	Van Meijgaard et al. (2012)
REMO2009	Climate Service Center	Hagemann (2002)	1	Jacob and Podzun (1997)