*Paper submitted to "The Cryosphere"*

# Snow water equivalent in the Alps as seen by gridded datasets and CMIP5 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)

**Reviewer: Yves Cornet, University of Liège (BE)**

**24 February 2017**

## Brief research description and my specific skills

The authors compare the snow water equivalent content provided by 35 Global Climatic Models and 4 Regional Climatic Models to 6 reference datasets (2 products derived from microwave remote sensing) over the Great Alpine Region for historical period (1980-2005). The lack of representative field observations and the coarse resolution of those products with regard to the high heterogeneity of mountain area are declared as limitations of the analysis whose results must be interpreted with caution.

Despite this statement, however, they exploit the different models to predict the future evolution until the middle of the 21st century of the annual cycle of snow cover in this region.

In the discussion section, they mention some specificities of the models that explain the discrepancies between their results.

As I don't know anything about these models I'm not able to evaluate the truthfulness of the discussion aspect regarding the concepts and data that govern the construction of these models. But as spatial analyst I will emphasis my comments on the research paradigms that are related to these skills.

My comments, remarks and criticisms are reported below and structured in 5 sections:

- Research paradigms and hypothesis to be demonstrated,
- Methodological issues,
- Specific scientific comments,
- Comments on the document's form (text, units, figures …),
- Conclusion.

## Research paradigms and hypothesis to be demonstrated

As recalled above, the authors are aware of the limitations of their research. The weaknesses of the product used are clearly recognized in the document. Even in this mountainous European area where a dense network of climatic stations is present in comparison to other parts of the World the number of field observations is too limited to represent fairly the spatial heterogeneity of the studied phenomenon. So the knowledge of the real wold is lacking to constrain and validate models. Field observation of snow cover can thus not be used as reference. 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.

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.

## Methodological issues

<u>Weighting procedure</u> 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.

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

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.

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?
- 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!!!

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!

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).

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.

## 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 !

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

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 ?

p.3 l. 14 What do you mean by <u>large scale</u>? 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.

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).

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

p. 10 l. 9 "… a wider <u>distribution</u> of SNW values, …" → "… a wider <u>statistical dispersion</u> of SNW values, …" (if I understand correctly)

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.

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!

p. 15 l. 3 "… wet precipitation bias …" Pleonasm! -→ "… overestimated precipitation …" or "… positive precipitation bias …"

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!

## 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 "$10^5$kg/$m^3$" 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!!!

## Conclusion

To conclude my review I consider that the paper is probably interesting for model developers as long as any prediction result is performed. So I recommend the editor to publish it with several major corrections as suggested above and the removal of section 4.3 and the related comments and discussions.